GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2005, 04:00:28 ; Search time 3259.53 Seconds
(without alignments)
2020.266 Million cell updates/sec

Title: US-10-657-740-1
Perfect score: 916
Sequence: 1 MDVTIQHPWFKRTLGPFYPS.....HAERAIPVSREEKPTSAPSS 173

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10657740/runat_27052005_165253_3329/app_query.fasta_1.590
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=150
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10657740_@CGN_1_1_5533_@runat_27052005_165253_3329 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	916	100.0	522	9	AY419529	AY419529 Homo sapi
2	916	100.0	536	2	BF726399	BF726399 by06a09.y
3	916	100.0	576	2	BF726236	BF726236 by03f01.y
4	916	100.0	578	4	BM721893	BM721893 UI-E-EO0-
5	916	100.0	592	2	BF727295	BF727295 by19e10.y
6	916	100.0	592	4	BM705926	BM705926 UI-E-DW0-
7	916	100.0	596	4	BM696581	BM696581 UI-E-DW0-
8	916	100.0	607	6	CD675250	CD675250 fs21c02.y
9	916	100.0	629	2	BF727028	BF727028 by15g05.y
10	916	100.0	630	4	BM696651	BM696651 UI-E-DW0-
11	916	100.0	633	6	CD672144	CD672144 fg10b07.y
12	916	100.0	659	4	BM706270	BM706270 UI-E-DW0-
13	916	100.0	681	5	BX118596	BX118596 BX118596
14	916	100.0	698	2	BF727324	BF727324 by19h12.y
15	916	100.0	724	4	BM722336	BM722336 UI-E-EO0-
16	912	99.6	559	4	BM722779	BM722779 UI-E-EO0-
17	911	99.5	635	4	BM697066	BM697066 UI-E-DW0-
18	910	99.3	532	2	BF726358	BF726358 by05d12.y
19	909	99.2	522	9	AY419530	AY419530 Pan trogl
20	907	99.0	577	2	BF726253	BF726253 by03h05.y
21	907	99.0	587	2	BF726422	BF726422 by06d05.y
22	906	98.9	597	2	BF726330	BF726330 by05b01.y
23	903	98.6	580	4	BM706139	BM706139 UI-E-DW0-
24	903	98.6	586	4	BM722650	BM722650 UI-E-EO0-
25	901	98.4	569	4	BM697368	BM697368 UI-E-DW0-
26	901	98.4	661	4	BM697160	BM697160 UI-E-DW0-
27	900	98.3	604	5	BQ640267	BQ640267 he26c10.y
28	898	98.0	577	2	BF727002	BF727002 by15c11.y
29	896	97.8	518	2	BF726854	BF726854 by13a10.y
30	893	97.5	523	2	BF726438	BF726438 by06f01.y
31	893	97.5	588	4	BM696799	BM696799 UI-E-DW0-
32	892	97.4	593	4	BM697101	BM697101 UI-E-DW0-
33	889	97.1	558	2	BF726679	BF726679 by10d02.y
34	888	96.9	558	4	BM696667	BM696667 UI-E-DW0-
35	887	96.8	508	4	BM696489	BM696489 UI-E-DW0-
36	887	96.8	663	4	BM686748	BM686748 UI-E-CQ0-
37	880	96.1	503	2	BF727205	BF727205 by18b11.y
38	879	96.0	519	4	BM696477	BM696477 UI-E-DW0-
39	877	95.7	536	4	BM686206	BM686206 UI-E-CI1-
40	876	95.6	542	5	BQ637175	BQ637175 he06f05.y
41	873	95.3	498	4	BM696853	BM696853 UI-E-DW0-
42	873	95.3	688	7	CF732528	CF732528 UI-M-HA0-
43	871	95.1	522	9	AY419531	AY419531 Mus muscu
44	871	95.1	552	7	CK628252	CK628252 ip09h08.y
45	871	95.1	557	7	CK627503	CK627503 io04g04.y
46	871	95.1	558	7	CK627631	CK627631 io06f03.y
47	871	95.1	566	7	CK627600	CK627600 io06c02.y
48	871	95.1	570	7	CK628219	CK628219 ip09d07.y
49	871	95.1	574	7	CK627249	CK627249 io01b06.y
50	871	95.1	579	7	CK628200	CK628200 ip09b11.y
51	871	95.1	582	7	CK628246	CK628246 ip09g11.y
52	871	95.1	582	7	CK628391	CK628391 ip11f11.y
53	871	95.1	586	7	CK627945	CK627945 ip04h08.y
54	871	95.1	590	7	CK627884	CK627884 ip04a08.y
55	871	95.1	590	7	CK628336	CK628336 ip10h08.y
56	871	95.1	590	7	CK628526	CK628526 ip13f05.y
57	871	95.1	594	7	CK628372	CK628372 ip11d12.y
58	871	95.1	601	7	CK627483	CK627483 io04d12.y
59	871	95.1	601	7	CK627500	CK627500 io04f12.y
60	871	95.1	606	6	CB849766	CB849766 MRA-1515
61	871	95.1	606	7	CK628546	CK628546 ip13h12.y
62	871	95.1	615	7	CK627740	CK627740 ip02b06.y
63	871	95.1	617	6	CB848065	CB848065 M2PN-3875
64	871	95.1	621	6	CB845999	CB845999 M2PN-1715
65	871	95.1	622	7	CK627736	CK627736 ip02b02.y
66	871	95.1	623	7	CK627777	CK627777 ip02f06.y
67	871	95.1	626	7	CK627695	CK627695 ip01e09.y
68	871	95.1	626	7	CK627737	CK627737 ip02b03.y
69	871	95.1	629	6	CB055541	CB055541 NISC_jj06
70	871	95.1	633	7	CK627814	CK627814 ip03b02.y
71	871	95.1	636	7	CK627616	CK627616 io06d08.y
72	871	95.1	639	7	CK627374	CK627374 io02h07.y
73	871	95.1	642	7	CK627318	CK627318 io02b04.y
74	871	95.1	643	6	CB846376	CB846376 M2PN-2110
75	871	95.1	643	7	CK627445	CK627445 io03h09.y
76	871	95.1	645	6	CB848071	CB848071 M2PN-3881
77	871	95.1	645	7	CK627651	CK627651 io06h10.y
78	871	95.1	662	6	CB841336	CB841336 M15E-1855
79	871	95.1	665	6	CB849833	CB849833 MRA-1582
80	871	95.1	669	6	CB840876	CB840876 M15E-1296
81	871	95.1	676	6	CB840784	CB840784 M15E-1183
82	871	95.1	690	6	CB846475	CB846475 M2PN-2213

83	871	95.1	691	7	CF731811	CF731811 UI-M-HA0-
84	871	95.1	696	7	CO429233	CO429233 UI-M-HW0-
85	871	95.1	699	7	CO427334	CO427334 UI-M-HW0-
86	871	95.1	705	7	CO429948	CO429948 UI-M-HW0-
87	871	95.1	709	7	CO428729	CO428729 UI-M-HW0-
88	871	95.1	721	7	CF733985	CF733985 UI-M-HA0-
89	871	95.1	851	6	CB845124	CB845124 M2PN-0694
90	869	94.9	677	7	CN435619	CN435619 BE030013B
91	867	94.7	600	4	BG805475	BG805475 0693-23 M
92	867	94.7	834	6	CB845009	CB845009 M2PN-0575
93	866	94.5	527	2	BF726880	BF726880 by13d10.y
94	866	94.5	536	2	BF726381	BF726381 by05g11.y
95	866	94.5	600	6	CA527708	CA527708 8051-64 M
96	866	94.5	637	6	CB848191	CB848191 M2PN-4009
97	866	94.5	854	6	CB848776	CB848776 MRA-0308
98	865	94.4	619	7	CK627238	CK627238 io01a02.y
99	865	94.4	625	7	CK627598	CK627598 io06b10.y
100	865	94.4	625	7	CO429950	CO429950 UI-M-HW0-
101	865	94.4	711	7	CO429062	CO429062 UI-M-HW0-
102	863	94.2	572	7	CK627552	CK627552 io05e04.y
103	863	94.2	600	4	BG808732	BG808732 2121-59 M
104	862	94.1	745	7	CN441369	CN441369 BE04025B2
105	862	94.1	809	7	CN438019	CN438019 BE04014A2
106	860	93.9	556	6	CB848450	CB848450 M2PN-4276
107	858	93.7	706	6	CB842697	CB842697 M15E-3361
108	857	93.6	645	6	CB842016	CB842016 M15E-2649
109	856	93.4	581	7	CF732553	CF732553 UI-M-HA0-
110	856	93.4	581	7	CO429470	CO429470 UI-M-HW0-
111	856	93.4	601	7	CO429510	CO429510 UI-M-HW0-
112	856	93.4	705	7	CO429072	CO429072 UI-M-HW0-
113	856	93.4	760	7	CO427914	CO427914 UI-M-HW0-
114	856	93.4	771	7	CO427433	CO427433 UI-M-HW0-
115	856	93.4	773	7	CO428121	CO428121 UI-M-HW0-
116	856	93.4	897	7	CR550671	CR550671 CR550671
117	854	93.2	506	2	BF727221	BF727221 by18e03.y
118	852	93.0	654	6	CB841950	CB841950 M15E-2579
119	851	92.9	653	6	CB850107	CB850107 MRA-1868
120	851	92.9	798	7	CO427114	CO427114 UI-M-HW0-
121	850	92.8	645	7	CO427082	CO427082 UI-M-HW0-
122	850	92.8	698	7	CN457234	CN457234 UI-M-HP0-
123	850	92.8	700	7	CO427294	CO427294 UI-M-HW0-
124	850	92.8	746	7	CF731540	CF731540 UI-M-HA0-
125	849.5	92.7	605	7	CK627795	CK627795 ip02h03.y
126	849.5	92.7	633	7	CO427584	CO427584 UI-M-HW0-
127	849.5	92.7	813	6	CB848822	CB848822 MRA-0355
128	849	92.7	485	2	BF727264	BF727264 by19b05.y
129	848	92.6	552	7	CK628267	CK628267 ip10b01.y
130	847	92.5	478	4	BM705752	BM705752 UI-E-DW0-
131	847	92.5	698	7	CO429320	CO429320 UI-M-HW0-
132	846	92.4	524	2	BF726441	BF726441 by06f04.y
133	845	92.2	539	7	CF732589	CF732589 UI-M-HA0-
134	845	92.2	580	7	CF731803	CF731803 UI-M-HA0-
135	845	92.2	650	6	CB842520	CB842520 M15E-3174
136	845	92.2	654	6	CB842155	CB842155 M15E-2792
137	845	92.2	769	7	CF732285	CF732285 UI-M-HA0-
138	844.5	92.2	778	7	CO427386	CO427386 UI-M-HW0-
139	844	92.1	759	7	CF731519	CF731519 UI-M-HA0-
140	842	91.9	522	2	BF726109	BF726109 by01b12.y
141	839	91.6	597	7	CF732732	CF732732 UI-M-HA0-
142	838	91.5	479	2	BF727402	BF727402 by20h11.y
143	838	91.5	519	2	BF726860	BF726860 by13b07.y
144	837	91.4	473	2	BF726931	BF726931 by14c03.y
145	837	91.4	585	6	CB848256	CB848256 M2PN-4074
146	837	91.4	642	4	BM721028	BM721028 UI-E-EO0-
147	836.5	91.3	686	7	CO427562	CO427562 UI-M-HW0-
148	835	91.2	516	2	BF727010	BF727010 by15d10.y
149	835	91.2	522	2	BF726595	BF726595 by09a09.y
150	835	91.2	560	7	CK627898	CK627898 ip04c07.y

ALIGNMENTS

RESULT 1

AY419529
LOCUS AY419529 522 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens CRYAA gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY419529
VERSION AY419529.1 GI:39775486
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 522)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 522)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .522
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1. .>522
/gene="CRYAA"
/locus_tag="HCM6921"
ORIGIN

Alignment Scores:

Pred. No.:	5.39e-101	Length:	522
Score:	916.00	Matches:	173
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-657-740-1 (1-173) x AY419529 (1-522)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 60

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 120

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 180

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 240

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 300

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 360
```

```
Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 420

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 480

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 519

RESULT 2
BF726399
LOCUS       BF726399                536 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by06a09.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by06a09 5', mRNA sequence.
ACCESSION   BF726399
VERSION     BF726399.1  GI:12042310
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 06  row: a  column: 09
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by06a09"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              5.59e-101       Length:         536
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0
```

```
US-10-657-740-1 (1-173) x BF726399 (1-536)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 71

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 131

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 191

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 251

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 311

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGNCGCTACCGCCTG 371

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 431

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 491

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        492 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 530

RESULT 3
BF726236
LOCUS       BF726236                576 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by03f01.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by03f01 5', mRNA sequence.
ACCESSION   BF726236
VERSION     BF726236.1  GI:12042147
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 03  row: f  column: 01
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by03f01"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
```

BY"
/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	6.17e-101	Length:	576
Score:	916.00	Matches:	173
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-657-740-1 (1-173) x BF726236 (1-576)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 107

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 167

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 227

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 287

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 347

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 407

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 467

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 527

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        528 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 566
```

RESULT 4
BM721893
LOCUS BM721893 578 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO0-ahw-c-15-0-UI.r1 UI-E-EO0 Homo sapiens cDNA clone UI-E-EO0-ahw-c-15-0-UI 5', mRNA sequence.
ACCESSION BM721893
VERSION BM721893.1 GI:19042101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 578)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EO0-ahw-c-15-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EO0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EO0 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:

Pred. No.:	6.2e-101	Length:	578
Score:	916.00	Matches:	173
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM721893 (1-578)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 114

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 174

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 234

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 294

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 354

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 414

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 474

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 534

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
                |||||||||||||||||||||||||||||||||||||
Db        535 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 573

RESULT 5
BF727295
LOCUS       BF727295                592 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by19e10.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by19e10 5', mRNA sequence.
ACCESSION   BF727295
VERSION     BF727295.1  GI:12043206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 19  row: e  column: 10
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by19e10"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:
Pred. No.:              6.41e-101       Length:         592
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-10-657-740-1 (1-173) x BF727295 (1-592)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 70

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 130

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 190

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 250

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 310

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 370

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 430

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 490

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
                |||||||||||||||||||||||||||||||||||||
Db        491 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 529

RESULT 6
BM705926
LOCUS       BM705926                592 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agh-f-18-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agh-f-18-0-UI 5', mRNA sequence.
ACCESSION   BM705926
VERSION     BM705926.1  GI:19019184
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
```

```
          Email: bento-soares@uiowa.edu
          Tissue Procurement: Dr. Gregg Hageman
           cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
           cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
           DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
           Clone Distribution: Researchers may obtain clones from Research
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agh-f-18-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                6.41e-101      Length:         592
Score:                    916.00         Matches:        173
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              100.00%        Indels:         0
DB:                       4              Gaps:           0

US-10-657-740-1 (1-173) x BM705926 (1-592)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 111

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 171

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 231

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 291

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 351

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 411

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 471

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 531

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        532 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 570

RESULT 7
BM696581
LOCUS       BM696581                596 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agk-o-15-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agk-o-15-0-UI 5', mRNA sequence.
ACCESSION   BM696581
VERSION     BM696581.1  GI:19009839
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agk-o-15-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                6.47e-101      Length:         596
Score:                    916.00         Matches:        173
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
```

```
Query Match:            100.00%      Indels:        0
DB:                     4            Gaps:          0

US-10-657-740-1 (1-173) x BM696581 (1-596)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 115

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 175

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 235

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 295

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 355

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 415

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 475

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 535

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        536 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 574

RESULT 8
CD675250
LOCUS       CD675250                607 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fs21c02.y1 Human Lens cDNA (Normalized): fs Homo sapiens cDNA clone
            fs21c02 5', mRNA sequence.
ACCESSION   CD675250
VERSION     CD675250.1  GI:32176981
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 607)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Behal,A., Touchman,J.W.,
            Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human lens for the NEIBank
            Project: over 2000 non-redundant transcripts, novel genes and
            splice variants
  JOURNAL   Mol. Vis. 8 (4), 171-184 (2002)
  MEDLINE   22103463
   PUBMED   12107413
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 21  row: c  column: 02
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .607
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fs21c02"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Normalized): fs"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human lens
                     library (by) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (COt 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              6.63e-101       Length:         607
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-10-657-740-1 (1-173) x CD675250 (1-607)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 110

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 170

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 230

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 290

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 350

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 410

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 470

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 530

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        531 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 569

RESULT 9
BF727028
LOCUS       BF727028                629 bp    mRNA    linear   EST 05-JAN-2001
```

```
DEFINITION  by15g05.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by15g05 5', mRNA sequence.
ACCESSION   BF727028
VERSION     BF727028.1  GI:12042939
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 15  row: g  column: 05
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .629
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by15g05"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                6.96e-101      Length:          629
Score:                    916.00         Matches:         173
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       2              Gaps:            0

US-10-657-740-1 (1-173) x BF727028 (1-629)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 107

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 167

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 227

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 287

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 347

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 407

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 467

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 527

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        528 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 566

RESULT 10
BM696651
LOCUS       BM696651                630 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agk-k-24-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agk-k-24-0-UI 5', mRNA sequence.
ACCESSION   BM696651
VERSION     BM696651.1  GI:19009909
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agk-k-24-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
```

I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:

Pred. No.:	6.98e-101	Length:	630
Score:	916.00	Matches:	173
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM696651 (1-630)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 113

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 173

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 233

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 293

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 353

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 413

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 473

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 533

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        534 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 572
```

RESULT 11
CD672144
LOCUS CD672144 633 bp mRNA linear EST 24-JUN-2003
DEFINITION fg10b07.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone fg10b07 5', mRNA sequence.
ACCESSION CD672144
VERSION CD672144.1 GI:32173875
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A., Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of adult human iris for the NEIBank Project: steroid-response factors and similarities with retinal pigment epithelium
JOURNAL Mol. Vis. 8 (4), 185-195 (2002)
MEDLINE 22103462
PUBMED 12107412
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 10 row: b column: 07
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fg10b07"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Iris cDNA (Normalized): fg"
/note="Organ: Eye; Vector: pCMVSPORT6; A human iris library (bx) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NotI. This NotI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (C0t 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center(NISC)."
ORIGIN

Alignment Scores:

Pred. No.:	7.02e-101	Length:	633
Score:	916.00	Matches:	173
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-657-740-1 (1-173) x CD672144 (1-633)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 111

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 171

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 231

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 291

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 351

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 411

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 471

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 531

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
                |||||||||||||||||||||||||||||||||||||
Db        532 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 570

RESULT 12
BM706270
LOCUS       BM706270                659 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agh-k-09-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agh-k-09-0-UI 5', mRNA sequence.
ACCESSION   BM706270
VERSION     BM706270.1  GI:19019528
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agh-k-09-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:              7.42e-101      Length:        659
Score:                  916.00         Matches:       173
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-10-657-740-1 (1-173) x BM706270 (1-659)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 116

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 176

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 236

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 296

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 356

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 416

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 476

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 536

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        537 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 575

RESULT 13
BX118596
LOCUS       BX118596                681 bp    mRNA    linear   EST 07-FEB-2003
DEFINITION  BX118596 Soares retina N2b4HR Homo sapiens cDNA clone
            IMAGp998P14436 ; IMAGE:220285, mRNA sequence.
ACCESSION   BX118596
VERSION     BX118596.1  GI:27882169
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 681)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998P14436.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
```

```
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .681
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGp998P14436 ; IMAGE:220285"
                      /sex="male"
                      /tissue_type="retina"
                      /dev_stage="55 year old"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares retina N2b4HR"
                      /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). The retinas were obtained from a 55 year old
                      Caucasian and total cellular poly(A)+ RNA was extracted 6
                      hrs after their removal. The retina RNA was kindly
                      provided by Roderick R. McInnes M.D. Ph.D. from the
                      University of Toronto. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. "
ORIGIN

Alignment Scores:
Pred. No.:              7.76e-101       Length:         681
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     5               Gaps:           0

US-10-657-740-1 (1-173) x BX118596 (1-681)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 152

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 212

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 272

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 332

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 392

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 452

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 512

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 572

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        573 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 611

RESULT 14
BF727324
LOCUS       BF727324                698 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by19h12.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by19h12 5', mRNA sequence.
ACCESSION   BF727324
VERSION     BF727324.1  GI:12043235
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 19  row: h  column: 12
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES              Location/Qualifiers
     source           1. .698
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="by19h12"
                      /tissue_type="Lens"
                      /dev_stage="Adult"
                      /lab_host="EMDH10B"
                      /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                      BY"
                      /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                      from different adults (both approximately 40 years old)
                      together yielded 20ug of total RNA and 150ng mRNA for cDNA
                      library synthesis. A directionally cloned cDNA library in
                      the pCMVSPORT6 vector was constructed at Life
                      Technologies, essentially following the protocols of the
                      SuperScript Plasmid System full details of which are
                      contained in the manufacturer's Instruction manual
                      (http://www.lifetech.com/). First strand synthesis was
                      carried out using a Not I primer-adapter
                      [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                      end inserts were cloned into the Not I/EcoR V sites in the
                      vector. EST analysis was performed on the unamplified
                      library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              8.03e-101       Length:         698
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-10-657-740-1 (1-173) x BF727324 (1-698)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 114

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 174
```

```
Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 234

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 294

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 354

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 414

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 474

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 534

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        535 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 573
```

RESULT 15
BM722336
LOCUS BM722336 724 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO0-ahx-i-06-0-UI.r1 UI-E-EO0 Homo sapiens cDNA clone UI-E-EO0-ahx-i-06-0-UI 5', mRNA sequence.
ACCESSION BM722336
VERSION BM722336.1 GI:19042842
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 724)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

```
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EO0-ahx-i-06-0-UI"
                     /tissue_type="fetal eye"
                     /dev_stage="fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EO0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EO0 is a cDNA library containing the following
                     tissue(s): fetal eye. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGCGTATACC. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.: 8.44e-101 Length: 724
Score: 916.00 Matches: 173
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1 (1-173) x BM722336 (1-724)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 110

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 170

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 230

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 290

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 350

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 410

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCNCTGTCTGCCGATGGCATGCTGACC 470

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 530

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        531 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 569
```

RESULT 16
BM722779
LOCUS BM722779 559 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO0-ahy-l-01-0-UI.r1 UI-E-EO0 Homo sapiens cDNA clone UI-E-EO0-ahy-l-01-0-UI 5', mRNA sequence.
ACCESSION BM722779
VERSION BM722779.1 GI:19043602
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EO0-ahy-l-01-0-UI"
                     /tissue_type="fetal eye"
                     /dev_stage="fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EO0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EO0 is a cDNA library containing the following
                     tissue(s): fetal eye. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGCGTATACC. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:              1.82e-100       Length:         559
Score:                  912.00          Matches:        172
Percent Similarity:     99.42%          Conservative:   0
Best Local Similarity:  99.42%          Mismatches:     1
Query Match:            99.56%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x BM722779 (1-559)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 95

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 155

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 215

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 275

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 335

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 395

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 455

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 515

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        516 GTGTCGCGGGAGGAGAAGCCCACCTCGGNTCCCTCGTCC 554

RESULT 17
BM697066
LOCUS       BM697066                635 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agm-g-03-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agm-g-03-0-UI 5', mRNA sequence.
ACCESSION   BM697066
VERSION     BM697066.1  GI:19010324
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agm-g-03-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
```

oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:

Pred. No.:	2.87e-100	Length:	635
Score:	911.00	Matches:	172
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.45%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM697066 (1-635)

```
Qy          2 AspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArg 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGG 60

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSer 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCG 120

Qy         42 SerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIle 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATC 180

Qy         62 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCC 240

Qy         82 ProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsn 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAAC 300

Qy        102 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCG 360

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTC 420

Qy        142 CysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTG 480

Qy        162 SerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||||||||||||||||||
Db        481 TCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 516
```

RESULT 18
BF726358
LOCUS BF726358 532 bp mRNA linear EST 05-JAN-2001
DEFINITION by05d12.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by05d12 5', mRNA sequence.
ACCESSION BF726358
VERSION BF726358.1 GI:12042269
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 532)
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 05 row: d column: 12
Seq primer: M13RP1 reverse primer (ABI).

FEATURES Location/Qualifiers
source 1. .532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="by05d12"
/tissue_type="Lens"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Lens cDNA (Un-normalized, unamplified): BY"
/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	2.98e-100	Length:	532
Score:	910.00	Matches:	172
Percent Similarity:	99.42%	Conservative:	0
Best Local Similarity:	99.42%	Mismatches:	1
Query Match:	99.34%	Indels:	0
DB:	2	Gaps:	0

US-10-657-740-1 (1-173) x BF726358 (1-532)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 61

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 121

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 181

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 241

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 301

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 361

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        362 CCGTCCAACGTGGACCAGTTGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 421

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        422 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 481

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        482 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 520

RESULT 19
AY419530
LOCUS       AY419530                522 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Pan troglodytes CRYAA gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY419530
VERSION     AY419530.1  GI:39775487
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 522)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>522
                     /gene="CRYAA"
                     /locus_tag="HCM6921"
ORIGIN

Alignment Scores:
Pred. No.:              3.84e-100      Length:         522
Score:                  909.00         Matches:        172
Percent Similarity:     99.42%         Conservative:   0
Best Local Similarity:  99.42%         Mismatches:     1
Query Match:            99.24%         Indels:         0
DB:                     9              Gaps:           0

US-10-657-740-1 (1-173) x AY419530 (1-522)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 60

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 120

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 180

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 240

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 300

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGAGCGCCAGNNNGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 360

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCCGCCGATGGCATGCTGACC 420

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 480

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 519

RESULT 20
BF726253
LOCUS       BF726253                577 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by03h05.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by03h05 5', mRNA sequence.
ACCESSION   BF726253
VERSION     BF726253.1  GI:12042164
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 03  row: h  column: 05
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by03h05"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
```

```
Pred. No.:              7.72e-100       Length:         577
Score:                  907.00          Matches:        172
Percent Similarity:     99.42%          Conservative:   0
Best Local Similarity:  99.42%          Mismatches:     1
Query Match:            99.02%          Indels:         0
DB:                     2               Gaps:           0

US-10-657-740-1 (1-173) x BF726253 (1-577)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 107

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 167

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 227

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 287

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 347

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 407

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 467

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCNC 527

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        528 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 566

RESULT 21
BF726422
LOCUS       BF726422                 587 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by06d05.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by06d05 5', mRNA sequence.
ACCESSION   BF726422
VERSION     BF726422.1  GI:12042333
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 587)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 06  row: d  column: 05
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .587
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by06d05"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              7.91e-100       Length:         587
Score:                  907.00          Matches:        172
Percent Similarity:     99.42%          Conservative:   0
Best Local Similarity:  99.42%          Mismatches:     1
Query Match:            99.02%          Indels:         0
DB:                     2               Gaps:           0

US-10-657-740-1 (1-173) x BF726422 (1-587)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 67

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 127

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 187

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 247

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 307

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 367

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 427

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCNC 487

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        488 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 526

RESULT 22
BF726330
LOCUS       BF726330                 597 bp    mRNA    linear   EST 05-JAN-2001
```

```
DEFINITION  by05b01.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by05b01 5', mRNA sequence.
ACCESSION   BF726330
VERSION     BF726330.1  GI:12042241
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 05  row: b  column: 01
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by05b01"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                1.07e-99        Length:          597
Score:                    906.00          Matches:         171
Percent Similarity:       98.84%          Conservative:    0
Best Local Similarity:    98.84%          Mismatches:      2
Query Match:              98.91%          Indels:          0
DB:                       2               Gaps:            0

US-10-657-740-1 (1-173) x BF726330 (1-597)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 61

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 121

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 181

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        182 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 241

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCCCCGGAGGACTTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 301

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 361

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGATC 421

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 481

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        482 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 520

RESULT 23
BM706139
LOCUS       BM706139                 580 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agg-b-16-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agg-b-16-0-UI 5', mRNA sequence.
ACCESSION   BM706139
VERSION     BM706139.1  GI:19019397
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agg-b-16-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
```

I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:

Pred. No.:	2.39e-99	Length:	580
Score:	903.00	Matches:	170
Percent Similarity:	99.42%	Conservative:	0
Best Local Similarity:	99.42%	Mismatches:	1
Query Match:	98.58%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM706139 (1-580)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 115

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 175

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 235

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 295

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 355

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 415

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCNCTGTCTGCCGATGGCATGCTGACC 475

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 TTCTGTGGCCCCAAGATCCAGACTGGGCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 535

Qy        161 ValSerArgGluGluLysProThrSerAlaPro 171
              |||||||||||||||||||    |||||||||
Db        536 GTGTCGCGGGAGGAGAAGCCCAACTCGGCTCCC 568
```

RESULT 24
BM722650
LOCUS BM722650 586 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO0-ahy-a-10-0-UI.r1 UI-E-EO0 Homo sapiens cDNA clone UI-E-EO0-ahy-a-10-0-UI 5', mRNA sequence.
ACCESSION BM722650
VERSION BM722650.1 GI:19043377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EO0-ahy-a-10-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EO0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EO0 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:

Pred. No.:	2.42e-99	Length:	586
Score:	903.00	Matches:	173
Percent Similarity:	99.43%	Conservative:	0
Best Local Similarity:	99.43%	Mismatches:	0
Query Match:	98.58%	Indels:	1
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM722650 (1-586)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 109

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 169

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 229

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 289

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 349

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 409

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 469

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThr-HisAlaGluArgAlaIlePr 160
                |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        470 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCCACGCCGAGCGAGCCATCCC 529

Qy        160 oValSerArgGluGluLysProThrSerAlaProSerSer 173
                |||||||||||||||||||||||||||||||||||||
Db        530 CGTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 569

RESULT 25
BM697368
LOCUS       BM697368                 569 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agm-d-06-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agm-d-06-0-UI 5', mRNA sequence.
ACCESSION   BM697368
VERSION     BM697368.1  GI:19010626
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 569)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .569
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agm-d-06-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                4.08e-99        Length:          569
Score:                    901.00          Matches:         170
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              98.36%          Indels:          0
DB:                       4               Gaps:            0

US-10-657-740-1 (1-173) x BM697368 (1-569)

Qy          4 ThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPhe 23
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTC 61

Qy         24 AspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThr 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACC 121

Qy         44 IleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGlu 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAG 181

Qy         64 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAG 241

Qy         84 AspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArg 103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGC 301

Qy        104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAAC 361

Qy        124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGNATGCTGACCTTCTGTGGC 421

Qy        144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGG 481

Qy        164 GluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||
Db        482 GAGGAGAAGCCCACCTCGGCTCCCTCGTCC 511

RESULT 26
BM697160
LOCUS       BM697160                 661 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agm-i-06-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agm-i-06-0-UI 5', mRNA sequence.
ACCESSION   BM697160
VERSION     BM697160.1  GI:19010418
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
```

```
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .661
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="UI-E-DW0-agm-i-06-0-UI"
                      /tissue_type="lens"
                      /dev_stage="adult"
                      /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                      /clone_lib="UI-E-DW0"
                      /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                      modified polylinker; Site_1: EcoR I; Site_2: Not I;
                      UI-E-DW0 is a cDNA library containing the following
                      tissue(s): lens. The library was constructed according to
                      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                      1996. First strand cDNA synthesis was primed with an
                      oligo-dT primer containing a Not I site. Double stranded
                      cDNA was ligated to an EcoR I adaptor, digested with Not
                      I, and cloned directionally into pT7T3-Pac vector. The
                      oligonucleotide used to prime the synthesis of
                      first-strand cDNA contains a library tag sequence that is
                      located between the Not I site and the (dT)18 tail. The
                      sequence tag for this library is CGATTAGCGA. This library
                      was created for the program, Gene Discovery in the Visual
                      System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                5.01e-99        Length:         661
Score:                    901.00          Matches:        170
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              98.36%          Indels:         0
DB:                       4               Gaps:           0

US-10-657-740-1 (1-173) x BM697160 (1-661)

Qy          4 ThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPhe 23
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTC 62

Qy         24 AspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThr 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACC 122

Qy         44 IleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGlu 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAG 182

Qy         64 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAG 242

Qy         84 AspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArg 103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGC 302

Qy        104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAAC 362

Qy        124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGC 422

Qy        144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGG 482

Qy        164 GluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||
Db        483 GAGGAGAAGCCCACCTCGGCTCCCTCGTCC 512

RESULT 27
BQ640267
LOCUS       BQ640267                 604 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  he26c10.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
            Homo sapiens cDNA clone he26c10 5', mRNA sequence.
ACCESSION   BQ640267
VERSION     BQ640267.1  GI:21764726
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of human retina for the NEIBank
            Project: Retbindin, an abundant, novel retinal cDNA and alternative
            splicing of other retina-preferred gene transcripts
  JOURNAL   Mol. Vis. 8 (4), 196-204 (2002)
  MEDLINE   22103461
   PUBMED   12107411
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 26  row: c  column: 10
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES              Location/Qualifiers
     source           1. .604
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="he26c10"
                      /tissue_type="Retina"
                      /dev_stage="Adult"
                      /lab_host="EMDH10B"
                      /clone_lib="Human Retina cDNA (Un-normalized,
                      unamplified): hd/he"
                      /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
                      was dissected from two 80 year old donors with no observed
                      eye disease. 100ug of total RNA was used for library
                      construction. A directionally cloned cDNA library in the
                      pSPORT1 vector (Life Technologies) was constructed at
                      Bioserve Biotechnology (Laurel MD) essentially following
                      the protocols of the SuperScript Plasmid System full
                      details of which are contained in the manufacturer's
                      Instruction manual (http://www.lifetech.com/). First
                      strand synthesis was carried out using a Not I
                      primer-adapter
                      [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis
                      was performed on the unamplified library at the NIH
                      Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                5.86e-99        Length:         604
Score:                    900.00          Matches:        170
Percent Similarity:       98.27%          Conservative:   0
Best Local Similarity:    98.27%          Mismatches:     3
Query Match:              98.25%          Indels:         0
DB:                       5               Gaps:           0
```

```
US-10-657-740-1 (1-173) x BQ640267 (1-604)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db         47 ATGGATGTGACCATCCAGCACCCCTGGTTGAAGCGCACCCTGGGGCCCTTCTACCCCAGC 106

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 166

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGT 226

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 TTCTTTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 286

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 346

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 406

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 466

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 526

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        527 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 565

RESULT 28
BF727002
LOCUS       BF727002                 577 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by15c11.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by15c11 5', mRNA sequence.
ACCESSION   BF727002
VERSION     BF727002.1  GI:12042913
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 15  row: c  column: 11
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by15c11"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              9.65e-99        Length:         577
Score:                  898.00          Matches:        170
Percent Similarity:     99.42%          Conservative:   0
Best Local Similarity:  99.42%          Mismatches:     1
Query Match:            98.03%          Indels:         0
DB:                     2               Gaps:           0

US-10-657-740-1 (1-173) x BF727002 (1-577)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeu 22
              ||||||||||||||||||||||||||||||||||    ||||||||||||||||||||||
Db          1 GTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGNGGCCCTTCTACCCCAGCCGGCTG 60

Qy         23 PheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSer 42
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCC 120

Qy         43 ThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSer 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCT 180

Qy         63 GluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerPro 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCG 240

Qy         83 GluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGlu 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAG 300

Qy        103 ArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSer 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCC 360

Qy        123 AsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCys 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGT 420

Qy        143 GlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSer 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCG 480

Qy        163 ArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||
Db        481 CGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 513

RESULT 29
BF726854
LOCUS       BF726854                 518 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by13a10.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by13a10 5', mRNA sequence.
ACCESSION   BF726854
VERSION     BF726854.1  GI:12042765
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 518)
  AUTHORS  Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE    NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL  Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT    Contact: Wistow G
           Section on Molecular Structure and Function
           National Eye Institute
           6/331, NIH, Bethesda, MD 20892-2740, USA
           Tel: 301 402 3452
           Fax: 301 496 0078
           Email: graeme@helix.nih.gov
           Plate: 13  row: a  column: 10
           Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by13a10"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                1.46e-98        Length:          518
Score:                    896.00          Matches:         169
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              97.82%          Indels:          0
DB:                       2               Gaps:            0

US-10-657-740-1 (1-173) x BF726854 (1-518)

Qy          5 IleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAsp 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTCGAC 61

Qy         25 GlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIle 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACCATC 121

Qy         45 SerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluVal 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTT 181

Qy         65 ArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CGATCCGACCGGGACAAGTTCGTCATCTTTCTCGATGTGAAGCACTTCTCCCCGGAGGAC 241

Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAG 301

Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTG 361

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCC 421

Qy        145 LysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGlu 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAG 481

Qy        165 GluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||
Db        482 GAGAAGCCCACCTCGGCTCCCTCGTCC 508

RESULT 30
BF726438
LOCUS       BF726438                523 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by06f01.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by06f01 5', mRNA sequence.
ACCESSION   BF726438
VERSION     BF726438.1  GI:12042349
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 523)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 06  row: f  column: 01
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by06f01"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                3.43e-98        Length:          523
Score:                    893.00          Matches:         168
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              97.49%          Indels:          0
```

```
DB:                 2           Gaps:          0

US-10-657-740-1 (1-173) x BF726438 (1-523)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 77

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 137

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 197

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 257

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 317

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 377

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 437

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 497

Qy        161 ValSerArgGluGluLysProThr 168
              |||||||||||||||||||||||
Db        498 GTGTCGCGGGAGGAGAAGCCCACC 521

RESULT 31
BM696799
LOCUS       BM696799                588 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agk-p-02-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agk-p-02-0-UI 5', mRNA sequence.
ACCESSION   BM696799
VERSION     BM696799.1  GI:19010057
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.

FEATURES             Location/Qualifiers
     source          1. .588
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agk-p-02-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                4.02e-98      Length:          588
Score:                    893.00        Matches:         171
Percent Similarity:       98.84%        Conservative:    0
Best Local Similarity:    98.84%        Mismatches:      2
Query Match:              97.49%        Indels:          1
DB:                       4             Gaps:            0

US-10-657-740-1 (1-173) x BM696799 (1-588)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 113

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 173

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 233

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 293

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 353

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 413

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 473

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 533

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||   |||   |||||||||
Db        534 GTGTCGCGGGAGGAGAAGCCACC-TCGGNTCCCTCGTCC 571
```

```
RESULT 32
BM697101
LOCUS       BM697101                593 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agm-m-13-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agm-m-13-0-UI 5', mRNA sequence.
ACCESSION   BM697101
VERSION     BM697101.1  GI:19010359
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .593
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agm-m-13-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                5.39e-98        Length:         593
Score:                    892.00          Matches:        168
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              97.38%          Indels:         0
DB:                       4               Gaps:           0

US-10-657-740-1 (1-173) x BM697101 (1-593)

Qy          6 GlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGln 25
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTCGACCAG 60

Qy         26 PhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSer 45
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACCATCAGC 120

Qy         46 ProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluValArg 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTTCGA 180

Qy         66 SerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeu 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGACCTC 240

Qy         86 ThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAsp 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAGGAC 300

Qy        106 AspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAsp 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTGGAC 360

Qy        126 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLys 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCAAG 420

Qy        146 IleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGlu 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAGGAG 480

Qy        166 LysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||
Db        481 AAGCCCACCTCGGCTCCCTCGTCC 504

RESULT 33
BF726679
LOCUS       BF726679                558 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by10d02.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by10d02 5', mRNA sequence.
ACCESSION   BF726679
VERSION     BF726679.1  GI:12042590
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 10  row: d  column: 02
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by10d02"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
```

the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	1.15e-97	Length:	558
Score:	889.00	Matches:	168
Percent Similarity:	99.41%	Conservative:	0
Best Local Similarity:	99.41%	Mismatches:	1
Query Match:	97.05%	Indels:	0
DB:	2	Gaps:	0

US-10-657-740-1 (1-173) x BF726679 (1-558)

```
Qy          5 IleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAsp 24
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db          2 ATCCAGCACCCCTGGTTCAAGCGCACCCTGNGGCCCTTCTACCCCAGCCGGCTGTTCGAC 61

Qy         25 GlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIle 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACCATC 121

Qy         45 SerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluVal 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTT 181

Qy         65 ArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGAC 241

Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAG 301

Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTG 361

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCC 421

Qy        145 LysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGlu 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AAGATCCAGACTGGCCTGGATGCAACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAG 481

Qy        165 GluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||||||||
Db        482 GAGAAGCCCACCTCGGCTCCCTCGTCC 508
```

RESULT 34
BM696667

LOCUS	BM696667 558 bp mRNA linear EST 28-FEB-2002
DEFINITION	UI-E-DW0-agk-o-10-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone UI-E-DW0-agk-o-10-0-UI 5', mRNA sequence.
ACCESSION	BM696667
VERSION	BM696667.1 GI:19009925
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 558)
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
PUBMED	8889548
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu Tissue Procurement: Dr. Gregg Hageman cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). Seq primer: M13 Reverse.
FEATURES	Location/Qualifiers
source	1. .558 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-E-DW0-agk-o-10-0-UI" /tissue_type="lens" /dev_stage="adult" /lab_host="DH10B (Life Technologies) (T1 phage resistant)" /clone_lib="UI-E-DW0" /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:

Pred. No.:	1.52e-97	Length:	558
Score:	888.00	Matches:	171
Percent Similarity:	98.84%	Conservative:	0
Best Local Similarity:	98.84%	Mismatches:	2
Query Match:	96.94%	Indels:	1
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM696667 (1-558)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 95

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 155

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 215

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 275

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        276 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 335

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 395

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        396 CCGTCCAACGTGGACCAGTCGGGCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 455

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 515

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||   ||||||||||||||||
Db        516 GTGTCGCGGGAGGAGAAGCCACC-TCGGCTCCCTCGTCC 553

RESULT 35
BM696489
LOCUS       BM696489                508 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-DW0-agj-n-20-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
            UI-E-DW0-agj-n-20-0-UI 5', mRNA sequence.
ACCESSION   BM696489
VERSION     BM696489.1  GI:19009747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .508
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agj-n-20-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                1.77e-97      Length:          508
Score:                    887.00        Matches:         167
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              96.83%        Indels:          0
DB:                       4             Gaps:            0

US-10-657-740-1 (1-173) x BM696489 (1-508)

Qy          7 HisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhe 26
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTCGACCAGTTT 60

Qy         27 PheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerPro 46
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACCATCAGCCCC 120

Qy         47 TyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSer 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTTCGATCC 180

Qy         67 AspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThr 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGACCTCACC 240

Qy         87 ValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAsp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAGGACGAC 300

Qy        107 HisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGln 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTGGACCAG 360

Qy        127 SerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIle 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCAAGATC 420

Qy        147 GlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLys 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAG 480

Qy        167 ProThrSerAlaProSerSer 173
              |||||||||||||||||||
Db        481 CCCACCTCGGCTCCCTCGTCC 501

RESULT 36
BM686748
LOCUS       BM686748                663 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-CQ0-adq-b-11-0-UI.r1 UI-E-CQ0 Homo sapiens cDNA clone
            UI-E-CQ0-adq-b-11-0-UI 5', mRNA sequence.
ACCESSION   BM686748
VERSION     BM686748.1  GI:19000006
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 663)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
```

```
           375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
           Tel: 319 335 8250
           Fax: 319 335 9565
           Email: bento-soares@uiowa.edu
           Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
            Clone Distribution: Researchers may obtain clones from Research
           Genetics (www.resgen.com).
           Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .663
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CQ0-adq-b-11-0-UI"
                     /tissue_type="optic nerve"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CQ0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CQ0 is a cDNA library containing the following
                     tissue(s): optic nerve. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CCATTAAGTG. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:                2.55e-97        Length:          663
Score:                    887.00          Matches:         171
Percent Similarity:       99.42%          Conservative:    0
Best Local Similarity:    99.42%          Mismatches:      1
Query Match:              96.83%          Indels:          1
DB:                       4               Gaps:            0

US-10-657-740-1 (1-173) x BM686748 (1-663)

Qy          2 AspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArg 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGG 69

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSer 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCG 129

Qy         42 SerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIle 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATC 189

Qy         62 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCC 249

Qy         82 ProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsn 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAAC 309

Qy        102 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCG 369

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTC 429

Qy        142 CysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 TGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTG 489

Qy        162 SerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||   ||||||||||||||||
Db        490 TCGCGGGAGGAGAAGCCACC-TCGGCTCCCTCGTCC 524

RESULT 37
BF727205
LOCUS       BF727205                503 bp    mRNA    linear   EST 05-JAN-2001
DEFINITION  by18b11.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
            sapiens cDNA clone by18b11 5', mRNA sequence.
ACCESSION   BF727205
VERSION     BF727205.1  GI:12043116
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  TITLE     NEIBANK: EST analysis and bioinformatics for ocular genomics
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 18  row: b  column: 11
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="by18b11"
                     /tissue_type="Lens"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
                     BY"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                     from different adults (both approximately 40 years old)
                     together yielded 20ug of total RNA and 150ng mRNA for cDNA
                     library synthesis. A directionally cloned cDNA library in
                     the pCMVSPORT6 vector was constructed at Life
                     Technologies, essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt
                     end inserts were cloned into the Not I/EcoR V sites in the
                     vector. EST analysis was performed on the unamplified
                     library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                1.25e-96        Length:          503
Score:                    880.00          Matches:         166
Percent Similarity:       99.40%          Conservative:    0
Best Local Similarity:    99.40%          Mismatches:      1
Query Match:              96.07%          Indels:          0
DB:                       2               Gaps:            0

US-10-657-740-1 (1-173) x BF727205 (1-503)
```

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGAGGCCCTTCTACCCCAGC 61

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 121

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 181

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 241

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCCCCAGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 301

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 361

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 421

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 481

Qy        161 ValSerArgGluGluLysPro 167
              |||||||||||||||||||||
Db        482 GTGTCGCGGGAGGAGAAGCCC 502
```

RESULT 38
BM696477
LOCUS BM696477 519 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-DW0-agj-1-20-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone UI-E-DW0-agj-1-20-0-UI 5', mRNA sequence.
ACCESSION BM696477
VERSION BM696477.1 GI:19009735
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 519)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

```
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-agj-1-20-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
ORIGIN
```

Alignment Scores:

Pred. No.:	1.72e-96	Length:	519
Score:	879.00	Matches:	166
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	1
Query Match:	95.96%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM696477 (1-519)

```
Qy          7 HisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhe 26
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 CACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTCGACCAGTTT 62

Qy         27 PheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerPro 46
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACCATCAGCCCC 122

Qy         47 TyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSer 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTTCGATCC 182

Qy         67 AspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThr 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGACCTCACC 242

Qy         87 ValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAsp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAGGACGAC 302

Qy        107 HisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGln 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTGGACCAG 362

Qy        127 SerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIle 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCAAGATC 422

Qy        147 GlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLys 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAA 482

Qy        167 ProThrSerAlaProSerSer 173
              |||||||||||||||||||
Db        483 GCCACCTCGGCTCCCTCGTCC 503
```

RESULT 39
BM686206
LOCUS BM686206 536 bp mRNA linear EST 27-FEB-2002
DEFINITION UI-E-CI1-aay-d-12-0-UI.r1 UI-E-CI1 Homo sapiens cDNA clone

```
          UI-E-CI1-aay-d-12-0-UI 5', mRNA sequence.
ACCESSION BM686206
VERSION   BM686206.1  GI:18996102
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1  (bases 1 to 536)
  AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE   Normalization and subtraction: two approaches to facilitate gene
          discovery
  JOURNAL Genome Res. 6 (9), 791-806 (1996)
  MEDLINE 97044477
   PUBMED 8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@uiowa.edu
          Tissue Procurement: Dr. Gregg Hageman
           cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
           cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
           DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
           Clone Distribution: Researchers may obtain clones from Research
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CI1-aay-d-12-0-UI"
                     /tissue_type="RPE and Choroid"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CI1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CI1 is a normalized cDNA library containing the
                     following tissue(s): RPE and Choroid. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is ACCTA.
                     This library was created for the program, Gene Discovery
                     in the Visual System, supported by National Eye Institute
                     (NEI)."
ORIGIN

Alignment Scores:
Pred. No.:              3.16e-96        Length:         536
Score:                  877.00          Matches:        165
Percent Similarity:     99.40%          Conservative:   0
Best Local Similarity:  99.40%          Mismatches:     1
Query Match:            95.74%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x BM686206 (1-536)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 98

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         99 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 158

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        159 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 218

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        219 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 278

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        279 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 338

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        339 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 398

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||

Db        399 CCGTCCAACGTGGACCAGTCGGGCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 458

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        459 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 518

Qy        161 ValSerArgGluGluLys 166
              ||||||||||||||||||
Db        519 GTGTCGCGGGAGGAGAAG 536

RESULT 40
BQ637175
LOCUS       BQ637175                542 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  he06f05.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
            Homo sapiens cDNA clone he06f05 5', mRNA sequence.
ACCESSION   BQ637175
VERSION     BQ637175.1  GI:21761634
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of human retina for the NEIBank
            Project: Retbindin, an abundant, novel retinal cDNA and alternative
            splicing of other retina-preferred gene transcripts
  JOURNAL   Mol. Vis. 8 (4), 196-204 (2002)
  MEDLINE   22103461
   PUBMED   12107411
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 06  row: f  column: 05
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="he06f05"
                     /tissue_type="Retina"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Retina cDNA (Un-normalized,
                     unamplified): hd/he"
                     /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
                     was dissected from two 80 year old donors with no observed
```

eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	4.24e-96	Length:	542
Score:	876.00	Matches:	165
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.63%	Indels:	0
DB:	5	Gaps:	0

US-10-657-740-1 (1-173) x BQ637175 (1-542)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 106

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 166

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 226

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 286

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 346

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 406

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 466

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 526

Qy        161 ValSerArgGluGlu 165
              |||||||||||||||
Db        527 GTGTCGCGGGAGGAG 541
```

RESULT 41
BM696853
LOCUS BM696853 498 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-DW0-agl-k-14-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone UI-E-DW0-agl-k-14-0-UI 5', mRNA sequence.
ACCESSION BM696853
VERSION BM696853.1 GI:19010111
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 498)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .498
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW0-agl-k-14-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:

Pred. No.:	8.76e-96	Length:	498
Score:	873.00	Matches:	164
Percent Similarity:	99.39%	Conservative:	0
Best Local Similarity:	99.39%	Mismatches:	1
Query Match:	95.31%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x BM696853 (1-498)

```
Qy          5 IleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAsp 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGCCGGCTGTTCGAC 61

Qy         25 GlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIle 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTGTCGTCCACCATC 121

Qy         45 SerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluVal 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTT 181

Qy         65 ArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGAC 241
```

```
Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAG 301

Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTG 361

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 144
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCAGTCGGNCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCC 421

Qy        145 LysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGlu 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAG 481

Qy        165 GluLysProThrSer 169
              |||||||||||||||
Db        482 GAGAAGCCCACCTCG 496
```

RESULT 42
CF732528

```
LOCUS       CF732528                688 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-HA0-cjw-j-03-0-UI.r1 NIH_BMAP_HA0 Mus musculus cDNA clone
            IMAGE:30551930 5', mRNA sequence.
ACCESSION   CF732528
VERSION     CF732528.1  GI:37628861
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 688)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .688
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30551930"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HA0"
                     /note="Organ: Eye; Vector: pYX- Asc;· Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin , Ph.D. ,
                     program coordinator."
```

ORIGIN

```
Alignment Scores:
Pred. No.:              1.36e-95        Length:         688
Score:                  873.00          Matches:        164
Percent Similarity:     97.69%          Conservative:   5
Best Local Similarity:  94.80%          Mismatches:     4
Query Match:            95.31%          Indels:         0
DB:                     7               Gaps:           0

US-10-657-740-1 (1-173) x CF732528 (1-688)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db         50 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 109

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 169

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 229

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 289

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||  :::||||||||||||||||||||||||||||
Db        290 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 349

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 409

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 469

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||  ||||||||:::|||:::|||||||||||||  |||:::||||||||||||||||
Db        470 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGTGCACAGCGAGAGGGCCATTCCT 529

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||:::|||||||||||||||
Db        530 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 568
```

RESULT 43
AY419531

```
LOCUS       AY419531                522 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Mus musculus CRYAA gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY419531
VERSION     AY419531.1  GI:39775488
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Clark,A.G., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 522)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
```

```
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>522
                     /gene="CRYAA"
                     /locus_tag="HCM6921"
ORIGIN

Alignment Scores:
Pred. No.:              1.64e-95        Length:         522
Score:                  871.00          Matches:        164
Percent Similarity:     97.69%          Conservative:   5
Best Local Similarity:  94.80%          Mismatches:     4
Query Match:            95.09%          Indels:         0
DB:                     9               Gaps:           0

US-10-657-740-1 (1-173) x AY419531 (1-522)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db          1 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 60

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 120

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 180

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 240

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        241 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 300

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 360

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 420

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   ||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        421 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 480

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||:::|||||||||||||||
Db        481 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 519

RESULT 44
CK628252
LOCUS       CK628252                552 bp    mRNA    linear   EST 26-JAN-2004
DEFINITION  ip09h08.y2 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA
            clone ip09h08 5', mRNA sequence.
ACCESSION   CK628252
VERSION     CK628252.1  GI:41349138
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Wistow,G. and Tomarev,S.
  TITLE     Expressed sequence tag analysis of mouse whole eye
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 09  row: h  column: 08
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .552
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57Bl6J"
                     /db_xref="taxon:10090"
                     /clone="ip09h08"
                     /tissue_type="Whole eye"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Mouse Whole eye, unamplified: io/ip"
                     /note="Organ: Eye; Vector: pSport1; Approximately 1mg
                     total RNA was extracted from 200 adult mouse whole eyes. A
                     directionally cloned cDNA library in the pSPORT1 vector
                     (Invitrogen) was constructed at Bioserve Biotechnology
                     (Laurel MD) essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was
                     cloned in Not I/Sal I sites. EST analysis was performed on
                     the unamplified library at the NIH Intramural Sequencing
                     Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              1.77e-95        Length:         552
Score:                  871.00          Matches:        164
Percent Similarity:     97.69%          Conservative:   5
Best Local Similarity:  94.80%          Mismatches:     4
Query Match:            95.09%          Indels:         0
DB:                     7               Gaps:           0

US-10-657-740-1 (1-173) x CK628252 (1-552)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db          5 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 64

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 124

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 184

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 244

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        245 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 304

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 364
```

```
Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 424

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        425 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 484

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||:::|||||||||||||||
Db        485 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 523

RESULT 45
CK627503
LOCUS       CK627503                557 bp    mRNA    linear   EST 26-JAN-2004
DEFINITION  io04g04.y1 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA
            clone io04g04 5', mRNA sequence.
ACCESSION   CK627503
VERSION     CK627503.1  GI:41348389
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Wistow,G. and Tomarev,S.
  TITLE     Expressed sequence tag analysis of mouse whole eye
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 04  row: g  column: 04
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57Bl6J"
                     /db_xref="taxon:10090"
                     /clone="io04g04"
                     /tissue_type="Whole eye"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Mouse Whole eye, unamplified: io/ip"
                     /note="Organ: Eye; Vector: pSport1; Approximately 1mg
                     total RNA was extracted from 200 adult mouse whole eyes. A
                     directionally cloned cDNA library in the pSPORT1 vector
                     (Invitrogen) was constructed at Bioserve Biotechnology
                     (Laurel MD) essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was
                     cloned in Not I/Sal I sites. EST analysis was performed on
                     the unamplified library at the NIH Intramural Sequencing
                     Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              1.79e-95        Length:         557
Score:                  871.00          Matches:        164
Percent Similarity:     97.69%          Conservative:   5
Best Local Similarity:  94.80%          Mismatches:     4
Query Match:            95.09%          Indels:         0
DB:                     7               Gaps:           0

US-10-657-740-1 (1-173) x CK627503 (1-557)
```

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db         11 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 70

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 130

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 190

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 250

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        251 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 310

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 370

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 430

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        431 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 490

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||:::|||||||||||||||
Db        491 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 529

RESULT 46
CK627631
LOCUS       CK627631                558 bp    mRNA    linear   EST 26-JAN-2004
DEFINITION  io06f03.y1 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA
            clone io06f03 5', mRNA sequence.
ACCESSION   CK627631
VERSION     CK627631.1  GI:41348517
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata;.Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Wistow,G. and Tomarev,S.
  TITLE     Expressed sequence tag analysis of mouse whole eye
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 06  row: f  column: 03
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57Bl6J"
                     /db_xref="taxon:10090"
                     /clone="io06f03"
                     /tissue_type="Whole eye"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Mouse Whole eye, unamplified: io/ip"
```

/note="Organ: Eye; Vector: pSport1; Approximately 1mg total RNA was extracted from 200 adult mouse whole eyes. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	1.79e-95	Length:	558
Score:	871.00	Matches:	164
Percent Similarity:	97.69%	Conservative:	5
Best Local Similarity:	94.80%	Mismatches:	4
Query Match:	95.09%	Indels:	0
DB:	7	Gaps:	0

US-10-657-740-1 (1-173) x CK627631 (1-558)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         11 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 70

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 130

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 190

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 250

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        251 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 310

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 370

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 430

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        431 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 490

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::|||||||||||||||
Db        491 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 529
```

RESULT 47
CK627600
LOCUS CK627600 566 bp mRNA linear EST 26-JAN-2004
DEFINITION io06c02.y1 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA clone io06c02 5', mRNA sequence.
ACCESSION CK627600
VERSION CK627600.1 GI:41348486
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 566)
AUTHORS Wistow,G. and Tomarev,S.
TITLE Expressed sequence tag analysis of mouse whole eye
JOURNAL Unpublished (2004)
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 06 row: c column: 02
Seq primer: M13RP1 reverse primer (ABI).

FEATURES Location/Qualifiers
source 1. .566
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl6J"
/db_xref="taxon:10090"
/clone="io06c02"
/tissue_type="Whole eye"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Mouse Whole eye, unamplified: io/ip"
/note="Organ: Eye; Vector: pSport1; Approximately 1mg total RNA was extracted from 200 adult mouse whole eyes. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	1.83e-95	Length:	566
Score:	871.00	Matches:	164
Percent Similarity:	97.69%	Conservative:	5
Best Local Similarity:	94.80%	Mismatches:	4
Query Match:	95.09%	Indels:	0
DB:	7	Gaps:	0

US-10-657-740-1 (1-173) x CK627600 (1-566)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         25 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 84

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 144

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 204

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 264

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        265 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 324

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 384
```

```
Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 444

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        445 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 504

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::|||||||||||||||
Db        505 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 543

RESULT 48
CK628219
LOCUS       CK628219                 570 bp    mRNA    linear   EST 26-JAN-2004
DEFINITION  ip09d07.y2 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA
            clone ip09d07 5', mRNA sequence.
ACCESSION   CK628219
VERSION     CK628219.1  GI:41349105
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Wistow,G. and Tomarev,S.
  TITLE     Expressed sequence tag analysis of mouse whole eye
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 09  row: d  column: 07
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57Bl6J"
                     /db_xref="taxon:10090"
                     /clone="ip09d07"
                     /tissue_type="Whole eye"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Mouse Whole eye, unamplified: io/ip"
                     /note="Organ: Eye; Vector: pSport1; Approximately 1mg
                     total RNA was extracted from 200 adult mouse whole eyes. A
                     directionally cloned cDNA library in the pSPORT1 vector
                     (Invitrogen) was constructed at Bioserve Biotechnology
                     (Laurel MD) essentially following the protocols of the
                     SuperScript Plasmid System full details of which are
                     contained in the manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was
                     cloned in Not I/Sal I sites. EST analysis was performed on
                     the unamplified library at the NIH Intramural Sequencing
                     Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              1.85e-95        Length:         570
Score:                  871.00          Matches:        164
Percent Similarity:     97.69%          Conservative:   5
Best Local Similarity:  94.80%          Mismatches:     4
Query Match:            95.09%          Indels:         0
DB:                     7               Gaps:           0

US-10-657-740-1 (1-173) x CK628219 (1-570)
```

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db         11 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 70

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 130

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 190

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 250

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||   :::||||||||||||||||||||||||||||
Db        251 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 310

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 370

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 430

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        431 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 490

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::|||||||||||||||
Db        491 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 529

RESULT 49
CK627249
LOCUS       CK627249                 574 bp    mRNA    linear   EST 26-JAN-2004
DEFINITION  io01b06.y1 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA
            clone io01b06 5', mRNA sequence.
ACCESSION   CK627249
VERSION     CK627249.1  GI:41348135
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Wistow,G. and Tomarev,S.
  TITLE     Expressed sequence tag analysis of mouse whole eye
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 01  row: b  column: 06
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .574
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57Bl6J"
                     /db_xref="taxon:10090"
                     /clone="io01b06"
                     /tissue_type="Whole eye"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Mouse Whole eye, unamplified: io/ip"
                     /note="Organ: Eye; Vector: pSport1; Approximately 1mg
```

total RNA was extracted from 200 adult mouse whole eyes. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	1.87e-95	Length:	574
Score:	871.00	Matches:	164
Percent Similarity:	97.69%	Conservative:	5
Best Local Similarity:	94.80%	Mismatches:	4
Query Match:	95.09%	Indels:	0
DB:	7	Gaps:	0

US-10-657-740-1 (1-173) x CK627249 (1-574)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         44 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 103

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 163

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 223

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 283

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              |||||||||||||||||||||||||   :::||||||||||||||||||||||||||||
Db        284 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 343

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 403

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 463

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        464 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 523

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::||||||||||||||||
Db        524 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 562
```

RESULT 50
CK628200
LOCUS CK628200 579 bp mRNA linear EST 26-JAN-2004
DEFINITION ip09b11.y2 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA clone ip09b11 5', mRNA sequence.
ACCESSION CK628200
VERSION CK628200.1 GI:41349086
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 579)
AUTHORS Wistow,G. and Tomarev,S.
TITLE Expressed sequence tag analysis of mouse whole eye
JOURNAL Unpublished (2004)
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 09 row: b column: 11
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .579
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl6J"
/db_xref="taxon:10090"
/clone="ip09b11"
/tissue_type="Whole eye"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Mouse Whole eye, unamplified: io/ip"
/note="Organ: Eye; Vector: pSport1; Approximately 1mg total RNA was extracted from 200 adult mouse whole eyes. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	1.89e-95	Length:	579
Score:	871.00	Matches:	164
Percent Similarity:	97.69%	Conservative:	5
Best Local Similarity:	94.80%	Mismatches:	4
Query Match:	95.09%	Indels:	0
DB:	7	Gaps:	0

US-10-657-740-1 (1-173) x CK628200 (1-579)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 114

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 174

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 234

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTGAAGCACTTC 294

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              |||||||||||||||||||||||||   :::||||||||||||||||||||||||||||
Db        295 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCACGGCAAACAC 354

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 414

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGCATGCTGACC 474

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
                |||   ||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        475 TTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGGGCCATTCCT 534

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
                ||||||||||||||||||||:::||||||||||||||||
Db        535 GTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 573
```

Search completed: May 30, 2005, 08:47:54
Job time : 3273.53 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 29, 2005, 23:18:18 ; Search time 4210.45 Seconds
(without alignments)
1990.942 Million cell updates/sec

Title: US-10-657-740-1
Perfect score: 916
Sequence: 1 MDVTIQHPWFKRTLGPFYPS.....HAERAIPVSREEKPTSAPSS 173

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10657740/runat_27052005_165253_3315/app_query.fasta_1.590
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=150
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10657740_@CGN_1_1_4362_@runat_27052005_165253_3315 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	916	100.0	519	9	CR407691	CR407691 Homo sapi
2	916	100.0	614	9	BC069528	BC069528 Homo sapi
3	916	100.0	741	9	HSU66584	U66584 Human alpha
4	916	100.0	1112	9	HSU05569	U05569 Human alpha
5	916	100.0	1114	6	CQ731849	CQ731849 Sequence
6	879	96.0	543	4	OCCRYAA	X95382 O.cuniculus
7	871	95.1	1056	10	RNU47922	U47922 Rattus norv
8	869	94.9	793	4	BOVCRYA	M26142 Bovine alph
9	849.5	92.7	796	10	MMU310308	AJ310308 Mus muscu
10	849.5	92.7	1271	10	RNU47921	U47921 Rattus norv
11	810	88.4	1025	10	MUSCRYAAB	J00376 Mouse lens
12	785	85.7	522	5	RCAACRYST	X85205 R.catesbeia
13	759	82.9	1361	5	D86299	D86299 Cynops pyrr
14	756	82.5	447	4	AJ617725	AJ617725 Elephas m
15	736	80.3	705	5	D88185	D88185 Xenopus lae
16	728.5	79.5	558	4	AJ617724	AJ617724 Ornithorh
17	715	78.1	477	5	AJ617726	AJ617726 Sphenodon
18	694	75.8	730	5	AY035778	AY035778 Danio rer
19	694	75.8	1430	5	BC083177	BC083177 Danio rer
20	677	73.9	448	5	TSU31938	U31938 Trachemys s
21	674	73.6	447	5	AJ617727	AJ617727 Lygodacty
22	663	72.4	448	5	EUDLAACRYS	L25850 Eudromia el
23	662	72.3	448	5	APRAAC	X96592 A.platyrhyn
24	653	71.3	448	5	CLRAAC	X96593 C.livia mRN
25	653	71.3	676	5	AY007972	AY007972 Clarias f
26	631	68.9	765	5	RTCRYA	X00716 Frog mRNA f
27	591	64.5	885	10	RNACRY	V01219 Partial seq
28	544	59.4	436	5	OLJ000940	AJ000940 Oryzias l
29	499	54.5	632	4	AF029793	AF029793 Bos tauru
30	493	53.8	548	4	OCCRYAB	X95383 O.cuniculus
31	489	53.4	528	9	BT006770	BT006770 Homo sapi
32	489	53.4	528	12	BT007909	BT007909 Synthetic
33	489	53.4	537	6	AX937702	AX937702 Sequence
34	489	53.4	537	6	AX937703	AX937703 Sequence
35	489	53.4	691	6	CQ812349	CQ812349 Sequence
36	489	53.4	691	6	AX330255	AX330255 Sequence
37	489	53.4	691	6	AX333032	AX333032 Sequence
38	489	53.4	691	6	BD079402	BD079402 Cancer-as
39	489	53.4	691	9	S45630	S45630 alpha B-cry
40	489	53.4	744	9	BC007008	BC007008 Homo sapi
41	489	53.4	856	6	AX888028	AX888028 Sequence
42	489	53.4	856	6	BD027638	BD027638 Sequence
43	489	53.4	893	6	AX899079	AX899079 Sequence
44	489	53.4	893	6	BD034612	BD034612 Sequence
45	489	53.4	911	6	AX899075	AX899075 Sequence
46	489	53.4	911	6	BD034608	BD034608 Sequence
47	489	53.4	927	6	AX899073	AX899073 Sequence
48	489	53.4	927	6	BD034606	BD034606 Sequence
49	489	53.4	942	6	AX888027	AX888027 Sequence
50	489	53.4	942	6	BD027637	BD027637 Sequence
c 51	489	53.4	1036	6	BD135185	BD135185 Human nuc
c 52	489	53.4	1036	6	AX017495	AX017495 Sequence
53	488	53.3	511	10	SEH272441	AJ272441 Spalax eh
54	488	53.3	528	10	SJU293658	AJ293658 Spalax ju
55	488	53.3	733	9	AB125159	AB125159 Macaca fa
56	485	52.9	661	9	AF007162	AF007162 Homo sapi
57	485	52.9	671	10	S74229	S74229 alpha B-cry
58	485	52.9	689	10	RRLENSABC	X60351 R.rattus mR
59	485	52.9	706	10	S77138	S77138 alpha B-cry
60	485	52.9	1247	6	AX401743	AX401743 Sequence
61	485	52.9	1247	10	RATCRYAB	M55534 Rat alpha-c
62	480	52.4	537	10	HAMSCARPB	J03849 Hamster alp
63	480	52.4	687	10	RRHARTABC	X60352 R.rattus mR
64	480	52.4	704	10	S77142	S77142 alpha B-cry
65	479	52.3	522	5	RCABCRYST	X87114 R.catesbeia
66	476	52.0	666	10	MUSCRYABA	M63170 Mouse alpha
67	476	52.0	848	10	BC010768	BC010768 Mus muscu
68	476	52.0	966	5	BC082122	BC082122 Xenopus l
69	470	51.3	507	4	AJ617732	AJ617732 Elephas m
70	458.5	50.1	900	5	AF159089	AF159089 Danio rer
71	457.5	49.9	913	6	AX899074	AX899074 Sequence
72	457.5	49.9	913	6	BD034607	BD034607 Sequence
73	454.5	49.6	466	4	AJ617731	AJ617731 Macropus
74	444.5	48.5	466	4	AJ617730	AJ617730 Didelphis
75	441.5	48.2	577	5	AY007973	AY007973 Clarias b
76	439.5	48.0	689	5	DUKABC	L08078 Anas platyr
77	437.5	47.8	846	5	GGU26661	U26661 Gallus gall

	78	437.5	47.8	1042	5	BX930014 BX930014 Gallus ga
	79	436.5	47.7	1251	5	S53164 S53164 alpha B-cry
	80	435.5	47.5	163962	9	BS000233 BS000233 Pan trogl
	81	435	47.5	148179	9	AP001631 AP001631 Homo sapi
	82	435	47.5	340000	9	AP001748 AP001748 Homo sapi
	83	430	46.9	888	5	BC075197 BC075197 Xenopus l
	84	427.5	46.7	501	4	AJ617728 AJ617728 Ornithorh
	85	407.5	44.5	502	4	AJ617729 AJ617729 Tachyglos
	86	406.5	44.4	826	6	AR415535 AR415535 Sequence
	87	406.5	44.4	826	6	AX972369 AX972369 Sequence
	88	406.5	44.4	826	6	BD111088 BD111088 EST and e
	89	404	44.1	249	10	RATAACRYA M96949 Rat alpha A
	90	401	43.8	249	10	RATAACRYB M96950 Rat alpha A
	91	400.5	43.7	6311	5	CHKCRYAA M17627 Chicken alp
	92	372.5	40.7	222895	2	AC084065 AC084065 Mus muscu
	93	370.5	40.4	1955	5	BC076518 BC076518 Danio rer
	94	367	40.1	419	9	AF026952 AF026952 Homo sapi
	95	366	40.0	167475	10	AC090881 AC090881 Mus Muscu
	96	349	38.1	1799	9	HSALPACRS X14789 H.sapiens a
	97	343.5	37.5	1065	5	AY184812 AY184812 Clarias b
	98	343	37.4	906	10	MAAACRY2 X02951 Hamster alp
c	99	343	37.4	166831	2	BX950201 BX950201 Danio rer
	100	342	37.3	2380	10	MMCRYS V00730 Mus musculu
	101	342	37.3	2403	10	MUSCRYAAA J00375 Mouse lens
	102	342	37.3	2580	10	MAAACRY1 X02950 Hamster alp
	103	341	37.2	1463	9	BC068046 BC068046 Homo sapi
	104	340	37.1	483	6	CQ721193 CQ721193 Sequence
	105	340	37.1	1457	9	AK056951 AK056951 Homo sapi
	106	340	37.1	1488	6	BD205158 BD205158 Human nuc
	107	340	37.1	1488	6	AX013767 AX013767 Sequence
	108	340	37.1	188810	2	AC121138 AC121138 Mus muscu
	109	337	36.8	165693	5	BX248514 BX248514 Zebrafish
	110	336	36.7	1310	10	D29960 D29960 Rattus norv
	111	335	36.6	189	9	HUMCRYAX1 M35629 Human alpha
	112	330	36.0	695	6	AX888029 AX888029 Sequence
	113	330	36.0	695	6	BD027639 BD027639 Sequence
	114	321.5	35.1	4321	5	AFAACEPIG Y11300 Astyanax fa
	115	321	35.0	1316	10	SEHCRYAA2 M17249 Mole rat al
	116	319.5	34.9	4339	5	AFAACCAVE Y11301 Astyanax fa
c	117	317.5	34.7	765	5	RTCRYA X00716 Frog mRNA f
	118	316	34.5	5491	10	SEHCRYAA1 M17247 Mole rat al
	119	308.5	33.7	719	6	CQ581247 CQ581247 Sequence
	120	308.5	33.7	735	3	AY047516 AY047516 Drosophil
	121	299	32.6	763	5	GGIAP X59541 Chicken mRN
	122	296.5	32.4	771	3	AF315318 AF315318 Bombyx mo
	123	294.5	32.2	686	3	AF315319 AF315319 Bombyx mo
	124	294.5	32.2	764	3	AF315317 AF315317 Bombyx mo
	125	293	32.0	168	9	HUMAACA L25781 Homo sapien
	126	293	32.0	767	3	PIU94328 U94328 Plodia inte
	127	292	31.9	535	10	MMU03562 U03562 Mus musculu
	128	291	31.8	349	6	AX786923 AX786923 Sequence
	129	290	31.7	830	3	AF237691 AF237691 Ciona int
	130	289.5	31.6	615	9	CR407614 CR407614 Homo sapi
	131	289.5	31.6	618	9	CR536489 CR536489 Homo sapi
	132	289.5	31.6	724	9	BC014920 BC014920 Homo sapi
	133	289.5	31.6	764	6	CQ799993 CQ799993 Sequence
	134	289.5	31.6	764	9	AB020027 AB020027 Homo sapi
	135	289.5	31.6	781	9	BC012768 BC012768 Homo sapi
	136	289.5	31.6	789	6	AR380755 AR380755 Sequence
	137	289.5	31.6	789	9	HSHSP27L X54079 Human mRNA
	138	289.5	31.6	794	9	BC073768 BC073768 Homo sapi
	139	289.5	31.6	847	6	BD186273 BD186273 STAT6 act
	140	289.5	31.6	865	6	CQ730135 CQ730135 Sequence
	141	289.5	31.6	865	9	HSU90906 U90906 Human clone
	142	289.5	31.6	867	9	BC000510 BC000510 Homo sapi
	143	289.5	31.6	1231	6	AX411221 AX411221 Sequence
	144	289.5	31.6	1231	9	HSHSP28 Z23090 H.sapiens m
	145	289.5	31.6	1380	6	AR217508 AR217508 Sequence
	146	289.5	31.6	1380	6	AX766880 AX766880 Sequence
	147	289	31.6	787	6	CQ858777 CQ858777 Sequence
	148	289	31.6	787	6	AX401752 AX401752 Sequence
	149	289	31.6	787	10	RATHSP27A M86389 Rattus norv
	150	288.5	31.5	797	3	DMEL2131 X77635 Drosophila

ALIGNMENTS

```
RESULT 1
CR407691
LOCUS       CR407691               519 bp    mRNA    linear   PRI 10-MAY-2004
DEFINITION  Homo sapiens full open reading frame cDNA clone RZPDo834E043D for
            gene CRYAA, crystallin, alpha A complete cds, without stopcodon.
ACCESSION   CR407691
VERSION     CR407691.1  GI:47115318
KEYWORDS    Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.
  TITLE     Cloning of human full open reading frames in Gateway(TM) system
            entry vector (pDONR201)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 519)
  AUTHORS   Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-2004) RZPD Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH Im Neuenheimer Feld 580, D-69120 Heidelberg,
            Germany
COMMENT     RZPD; RZPDo834E043D, ORFNo 642
            www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDo834E043D RZPDLIB;
            Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
            834
            www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
            www.rzpd.de/products/orfclones/
            Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            contact RZPD (customer.service@rzpd.de) for further information.
            This CDS clone is a part of a collection of human full length
            expression clones generated by RZPD.
            This CDS has been cloned without stopcodon.
            The CDS has been inserted into pDONR201 via a BP Clonase(TM)
            reaction. Additional sequence has been added in front of the start
            codon (ATG): att. .AAAAAA GCT GGC ACC CCT GGT CCA GGT (ATG)
            After the last codon additional sequence has been added: CCA GGC
            CCA GGC GGC G in front of the 3' att site (AC CCA GCT TTC TT).
            Compared to the reference sequence U66584 we did not find any amino
            acid exchanges.
            Clone distribution: http://www.rzpd.de/products/orfclones/.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDo834E043D"
                     /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
                     /lab_host="DH10B"
                     /note="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
     gene            1. .519
                     /gene="CRYAA"
     CDS             1. .>519
                     /gene="CRYAA"
                     /codon_start=1
                     /protein_id="CAG28619.1"
                     /db_xref="GI:47115319"
                     /translation="MDVTIQHPWFKRTLGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIP
                     VSREEKPTSAPSS"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                9.38e-94        Length:         519
Score:                    916.00          Matches:        173
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-10-657-740-1 (1-173) x CR407691 (1-519)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 60

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 120

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 180

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 240

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 300

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 360

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 420

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 480

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        481 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 519

RESULT 2
BC069528
LOCUS       BC069528                 614 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens crystallin, alpha A, mRNA (cDNA clone MGC:96924
            IMAGE:7262133), complete cds.
ACCESSION   BC069528
VERSION     BC069528.1  GI:46854598
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 614)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 614)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Baylor Human Genome Sequencing Center
            cDNA Library Preparation: Baylor Human Genome Sequencing Center
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRBR Plate: 1 Row: d Column: 7.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:96924 IMAGE:7262133"
                     /tissue_type="PCR rescued clones"
                     /clone_lib="NIH_MGC_244"
                     /note="Vector: pPCR-Script Amp SK(+)"
     gene            1. .614
                     /gene="CRYAA"
                     /note="synonyms: CRYA1, HSPB4"
                     /db_xref="LocusID:1409"
                     /db_xref="MIM:123580"
     CDS             51. .572
                     /gene="CRYAA"
                     /codon_start=1
                     /product="crystallin, alpha A"
                     /protein_id="AAH69528.1"
                     /db_xref="GI:46854599"
                     /db_xref="LocusID:1409"
                     /db_xref="MIM:123580"
                     /translation="MDVTIQHPWFKRTLGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIP
                     VSREEKPTSAPSS"
ORIGIN

Alignment Scores:
Pred. No.:                1.16e-93        Length:         614
Score:                    916.00          Matches:        173
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-10-657-740-1 (1-173) x BC069528 (1-614)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
```

```
Db         51 ATGGATGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 110

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40

Db        111 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 170

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60

Db        171 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 230

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80

Db        231 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 290

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100

Db        291 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 350

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120

Db        351 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 410

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140

Db        411 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 470

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160

Db        471 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 530

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173

Db        531 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 569
```

```
RESULT 3
HSU66584
LOCUS       HSU66584                741 bp    mRNA    linear   PRI 14-DEC-1996
DEFINITION  Human alphaA-crystallin (CRYAA) mRNA, complete cds.
ACCESSION   U66584
VERSION     U66584.1  GI:1732062
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Petrash,J.M., Mathur,S., Manoharan,M. and Andley,U.P.
  TITLE     Cloning and expression of human lens crystallins
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 36, S882-S882 (1995)
REFERENCE   2  (bases 1 to 741)
  AUTHORS   Andley,U.P., Mathur,S., Griest,T.A. and Petrash,J.M.
  TITLE     Cloning, expression, and chaperone-like activity of human
            alphaA-crystallin
  JOURNAL   J. Biol. Chem. 271 (50), 31973-31980 (1996)
  MEDLINE   97112991
   PUBMED   8943244
REFERENCE   3  (bases 1 to 741)
  AUTHORS   Petrash,J.M., Mathur,S., Wang,J.C., Griest,T.A. and Andley,U.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-1996) Ophthalmology and Visual Sciences,
            Washington University School of Medicine, 660 S. Euclid Ave., St.
            Louis, MO 63110, USA
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="eye lens"
     gene            1. .741
                     /gene="CRYAA"
     CDS             1. .522
                     /gene="CRYAA"
                     /function="structural protein"
                     /note="structurally related to sHSPs; chaperone-like
                     activity"
                     /codon_start=1
                     /product="alphaA-crystallin"
                     /protein_id="AAC50900.1"
                     /db_xref="GI:1732063"
                     /translation="MDVTIQHPWFKRTLGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIP
                     VSREEKPTSAPSS"
ORIGIN

Alignment Scores:
Pred. No.:                1.47e-93        Length:          741
Score:                    916.00          Matches:         173
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-10-657-740-1 (1-173) x HSU66584 (1-741)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20

Db          1 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 60

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40

Db         61 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 120

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60

Db        121 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 180

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80

Db        181 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 240

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100

Db        241 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 300

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120

Db        301 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 360

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140

Db        361 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 420

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160

Db        421 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 480

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173

Db        481 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 519
```

```
RESULT 4
HSU05569
LOCUS       HSU05569               1112 bp    mRNA    linear   PRI 25-APR-1996
DEFINITION  Human alphaA-crystallin (CRYA1) mRNA, complete cds.
ACCESSION   U05569
VERSION     U05569.1  GI:452477
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1112)
  AUTHORS   Jaworski,C.J.
  TITLE     The human alphaA-crystallin gene
```

```
  JOURNAL   Thesis (1992) LMDB, NEI, Molecular Structure and Function
REFERENCE   2  (bases 1 to 1112)
  AUTHORS   Jaworski,C.J.
  TITLE     A reassessment of mammalian alpha A-crystallin sequences using DNA
            sequencing: implications for anthropoid affinities of tarsier
  JOURNAL   J. Mol. Evol. 41 (6), 901-908 (1995)
  MEDLINE   96139023
   PUBMED   8587135
REFERENCE   3  (bases 1 to 1112)
  AUTHORS   Wistow,G.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-1994) Graeme J. Wistow, Molecular Structure and
            Function, LMDB, NEI, NIH, Bethesda, MD 20892, USA
FEATURES             Location/Qualifiers
     source          1. .1112
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /tissue_type="lens"
     gene            1. .1112
                     /gene="CRYA1"
     CDS             68. .589
                     /gene="CRYA1"
                     /function="lens structural protein"
                     /note="This sequence is derived from genomic and cDNA
                     sequence"
                     /citation=[1]
                     /citation=[2]
                     /codon_start=1
                     /evidence=experimental
                     /product="alphaA-crystallin"
                     /protein_id="AAA97523.1"
                     /db_xref="GI:452478"
                     /translation="MDVTIQHPWFKRTLGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIP
                     VSREEKPTSAPSS"
ORIGIN

Alignment Scores:
Pred. No.:              2.47e-93        Length:         1112
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-10-657-740-1 (1-173) x HSU05569 (1-1112)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 127

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 187

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 247

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 307

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 367

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 427

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 487

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 547

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        548 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 586

RESULT 5
CQ731849
LOCUS       CQ731849                1114 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 17783 from Patent WO02068579.
ACCESSION   CQ731849
VERSION     CQ731849.1  GI:42309516
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 17783 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1114
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.47e-93        Length:         1114
Score:                  916.00          Matches:        173
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-10-657-740-1 (1-173) x CQ731849 (1-1114)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 129

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 189

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 249

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 309

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 369

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 429

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 489

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 549

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
                |||||||||||||||||||||||||||||||||||||
Db        550 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 588

RESULT 6
OCCRYAA
LOCUS       OCCRYAA                   543 bp    mRNA    linear   MAM 23-OCT-1996
DEFINITION  O.cuniculus mRNA for alpha-A-crystallin.
ACCESSION   X95382
VERSION     X95382.1  GI:1177576
KEYWORDS    alpha-A-crystallin; cryaA gene.
SOURCE      Oryctolagus cuniculus (rabbit)
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE   1
  AUTHORS   Krausz,E., Augusteyn,R.C., Quinlan,R.A., Reddan,J.R., Russell,P.,
            Sax,C.M. and Graw,J.
  TITLE     Expression of Crystallins, Pax6, Filensin, CP49, MIP, and MP20 in
            lens-derived cell lines
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 37 (10), 2120-2128 (1996)
  MEDLINE   96409169
   PUBMED   8814151
REFERENCE   2  (bases 1 to 543)
  AUTHORS   Graw,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-1996) J. Graw, Institute of Mammalian Genetics,
            GSF-Research Center Neuherberg, Ingolstaedter Landstrasse 1,
            Oberschleissheim, D-85764, FRG
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Oryctolagus cuniculus"
                     /mol_type="mRNA"
                     /sub_species="New Zeeland white"
                     /db_xref="taxon:9986"
                     /tissue_type="lens"
     gene            1. .543
                     /gene="cryaa"
     CDS             22. .543
                     /gene="cryaa"
                     /codon_start=1
                     /product="alpha-A-crystallin"
                     /protein_id="CAA64668.1"
                     /db_xref="GI:1177577"
                     /db_xref="GOA:P02493"
                     /db_xref="UniProt/Swiss-Prot:P02493"
                     /translation="MDVTIQHPWFKRTLGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQEDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFSGPKVQSGLDAGHSERAIP
                     VSREEKPSSVPSS"
ORIGIN

Alignment Scores:
Pred. No.:              1.47e-89        Length:         543
Score:                  879.00          Matches:        165
Percent Similarity:     98.27%          Conservative:   5
Best Local Similarity:  95.38%          Mismatches:     3
Query Match:            95.96%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x OCCRYAA (1-543)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGGACGTCACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 81

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 CGGCTGTTCGACCAGTTCTTCGGTGAGGGCCTCTTCGAGTATGACCTGCTGCCCTTCCTG 141

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 201

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 ATCTCTGAGGTGCGGTCCGACCGGGACAAGTTCGTCATCTTCCTGGACGTGAAGCACTTC 261

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        262 TCGCCCGAGGACCTCACCGTGAAGGTGCAGGAGGACTTCGTGGAGATCCACGGCAAACAC 321

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AACGAGAGACAGGATGACCACGGCTACATTTCCCGCGAGTTCCACCGGCGCTACCGCCTG 381

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CCTTCCAACGTGGACCAGTCGGCGCTGTCCTGCTCGCTCTCCGCGGACGGCATGCTCACC 441

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::|||||||||||||    |||:::|||||||||||||||||
Db        442 TTCTCCGGCCCCAAGGTGCAGTCCGGCCTGGACGCTGGCCACAGCGAGCGAGCCATCCCC 501

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||:::|||    |||||||||
Db        502 GTGTCCCGGGAGGAGAAGCCCAGCTCGGTGCCCTCGTCC 540

RESULT 7
RNU47922
LOCUS       RNU47922                1056 bp    mRNA    linear   ROD 02-APR-1996
DEFINITION  Rattus norvegicus alpha A-crystallin mRNA, complete cds.
ACCESSION   U47922
VERSION     U47922.1  GI:1245161
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1056)
  AUTHORS   Bhat,S.P., Nandy,P., Srinivasan,A., Cheng,D. and Sitay,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-1996) Suraj P. Bhat, JSEI, UCLA School Medicine,
            100-Stein Plaza, Rm. BH623, Los Angeles, CA 90095-7008, USA
FEATURES             Location/Qualifiers
     source          1. .1056
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS             13. .534
                     /note="similar to small heat shock proteins"
                     /codon_start=1
                     /product="alpha A-crystallin"
                     /protein_id="AAA93367.1"
                     /db_xref="GI:1245162"
                     /translation="MDVTIQHPWFKRALGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVLEDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFSGPKVQSGLDAGHSERAIP
                     VSREEKPSSAPSS"
     3'UTR           535. .1050
ORIGIN

Alignment Scores:
Pred. No.:              2.73e-88        Length:         1056
Score:                  871.00          Matches:        164
Percent Similarity:     97.69%          Conservative:   5
Best Local Similarity:  94.80%          Mismatches:     4
```

```
Query Match:            95.09%      Indels:        0
DB:                     10          Gaps:          0

US-10-657-740-1 (1-173) x RNU47922 (1-1056)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 72

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 132

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 192

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTGAAGCACTTC 252

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||||
Db        253 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCATGGCAAACAC 312

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 372

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGCATGCTGACC 432

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::|||||||||||   |||:::||||||||||||||||
Db        433 TTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGGGCCATTCCC 492

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::||||||||||||||||
Db        493 GTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 531

RESULT 8
BOVCRYA
LOCUS       BOVCRYA                  793 bp    mRNA    linear   MAM 26-APR-1993
DEFINITION  Bovine alpha-A-crystallin gene, complete cds.
ACCESSION   M26142
VERSION     M26142.1  GI:162909
KEYWORDS    crystallin.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1  (bases 1 to 793)
  AUTHORS   Hay,R.E. and Petrash,J.M.
  TITLE     Nucleotide sequence of a bovine lens alpha A-crystallin cDNA
  JOURNAL   Biochem. Biophys. Res. Commun. 148 (1), 31-37 (1987)
  MEDLINE   88049675
   PUBMED   3675580
COMMENT     Original source text: Bovine lens, cDNA to mRNA, clone
            pBL-alpha-A2-1.
FEATURES             Location/Qualifiers
     source          1. .793
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
     CDS             31. .552
                     /note="alpha-A-crystallin"
                     /codon_start=1
                     /protein_id="AAA30471.1"
                     /db_xref="GI:162910"
                     /translation="MDIAIQHPWFKRTLGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQEDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFSGPKIPSGVDAGHSERAIP
                     VSREEKPSSAPSS"
ORIGIN      29 bp upstream of NcoI site.

Alignment Scores:
Pred. No.:              3.18e-88      Length:        793
Score:                  869.00        Matches:       163
Percent Similarity:     97.69%        Conservative:  6
Best Local Similarity:  94.22%        Mismatches:    4
Query Match:            94.87%        Indels:        0
DB:                     4             Gaps:          0

US-10-657-740-1 (1-173) x BOVCRYA (1-793)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||:::   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ATGGATATCGCCATTCAGCACCCCTGGTTCAAACGCACCCTGGGCCCCTTCTACCCCAGC 90

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 CGGCTGTTCGACCAGTTCTTCGGCGAGGGCCTCTTCGAGTACGACCTGCTGCCCTTCCTG 150

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 TCCTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 210

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ATCTCTGAGGTCCGATCCGACCGGGACAAGTTTGTCATCTTCCTGGATGTGAAGCACTTC 270

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        271 TCTCCCGAGGACCTGACGGTGAAGGTGCAGGAGGACTTCGTGGAGATCCACGGCAAGCAC 330

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 AACGAGCGGCAGGATGACCATGGCTACATCTCCCGCGAGTTCCACCGCCGCTACCGCCTG 390

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CCTTCCAACGTGGACCAGTCCGCACTCTCCTGCTCCCTGTCCGCTGATGGCATGCTGACC 450

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||||   :::|||:::|||||||   |||:::||||||||||||||||
Db        451 TTCTCTGGCCCCAAGATCCCATCTGGCGTGGACGCCGGCCACAGCGAGCGGGCCATCCCC 510

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::||||||||||||||||
Db        511 GTGTCCCGGGAGGAGAAGCCCAGCTCTGCGCCCTCGTCC 549

RESULT 9
MMU310308
LOCUS       MMU310308                796 bp    mRNA    linear   ROD 03-APR-2001
DEFINITION  Mus musculus mRNA for alpha-A-crystallin (Cryaa gene).
ACCESSION   AJ310308
VERSION     AJ310308.1  GI:13548627
KEYWORDS    alpha-A-crystallin; cryaA gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Graw,J.
  TITLE     Characterization of a new, dominant V124E mutation in the mouse
            alpha-A-crystallin encoding gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 796)
  AUTHORS   Graw,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2001) Graw J., Institute of Mammalian Genetics,
            GSF-National Research Center for Environment and Health,
            Ingolstaedter Landstr. 1, Neuherberg, Germany, D-85764, GERMANY
FEATURES             Location/Qualifiers
```

```
     source          1. .796
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C3H/El"
                     /db_xref="taxon:10090"
                     /tissue_type="Lens"
     gene            1. .796
                     /gene="Cryaa"
     CDS             52. .642
                     /gene="Cryaa"
                     /function="lens structural protein; chaperone"
                     /note="alternative splicing"
                     /codon_start=1
                     /product="alpha-A-crystallin"
                     /protein_id="CAC35974.1"
                     /db_xref="GI:13548628"
                     /translation="MDVTIQHPWFKRALGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISELMTHMWFVMHQPHAGNPKNNPVKVRSDRDKFVIFLDVKH
                     FSPEDLTVKVLEDFVEIHGKHNERQDDHGYISREFHRRYRLPSNVDQSALSCSLSADG
                     MLTFSGPKVQSGLDAGHSERAIPVSREEKPSSAPSS"
     exon            241. .309
                     /gene="Cryaa"
                     /note="alternative splicing"
ORIGIN

Alignment Scores:
Pred. No.:              5.04e-86        Length:         796
Score:                  849.50          Matches:        164
Percent Similarity:     86.22%          Conservative:   5
Best Local Similarity:  83.67%          Mismatches:     4
Query Match:            92.74%          Indels:         23
DB:                     10              Gaps:           1

US-10-657-740-1 (1-173) x MMU310308 (1-796)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         52 ATGGACGTCACCATTCAGCATCCTTGGTTCAAGCGTGCCCTGGGGCCCTTCTACCCCAGC 111

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAGTACGACCTGCTGCCCTTCCTG 171

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TCTTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACTGTGCTGGACTCGGGC 231

Qy         61 IleSerGlu----------------------------------------------------- 63
              |||||||||
Db        232 ATCTCTGAGCTCATGACCCATATGTGGTTTGTAATGCACCAACCACATGCTGGAAACCCC 291

Qy         64 ------------------ValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                                 ||||||||||||||||||||||||||||||||||||||||
Db        292 AAGAACAACCCCGTCAAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGACGTG 351

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||||||||||||||||||||||   :::||||||||||||||||||
Db        352 AAGCACTTCTCTCCTGAGGACCTCACCGTGAAGGTACTGGAGGATTTTGTGGAGATTCAC 411

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 GGCAAACACAACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGC 471

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TACCGTCTGCCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCCTGTCTGCGGATGGC 531

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              ||||||||||||   |||||||||:::|||:::||||||||||||   |||:::||||||
Db        532 ATGCTGACCTTCTCTGGCCCCAAGGTCCAGTCCGGTTTGGATGCTGGCCACAGCGAGAGG 591

Qy        158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||||||||||||:::||||||||||||||||
```

```
Db        592 GCCATTCCTGTGTCACGGGAGGAGAAACCCAGCTCTGCACCCTCGTCC 639

RESULT 10
RNU47921
LOCUS       RNU47921                1271 bp    mRNA    linear   ROD 02-APR-1996
DEFINITION  Rattus norvegicus alpha A (insert)-crystallin mRNA, complete cds.
ACCESSION   U47921
VERSION     U47921.1  GI:1245159
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1271)
  AUTHORS   Bhat,S.P., Nandy,P., Srinivasan,A., Cheng,D. and Sitay,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-1996) Suraj P. Bhat, JSEI, UCLA School Medicine,
            100-Stein Plaza, Rm. BH623, Los Angeles, CA 90095-7008, USA
FEATURES             Location/Qualifiers
     source          1. .1271
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="lens"
     CDS             159. .749
                     /note="alpha A (insert)-crystallin is similar to alpha
                     A-crystallin, the predominant protein of the ocular lens.
                     These two proteins are identical in sequence except for
                     the presence of a 23 amino acid peptide in alpha
                     A(insert)-crystallin located between the 63rd and the 64th
                     residue of alpha A-crystallin; alpha A(insert)-crystallin
                     represents about 10-20% of total alpha crystallin in the
                     lens"
                     /codon_start=1
                     /product="alpha A (insert)-crystallin"
                     /protein_id="AAA93366.1"
                     /db_xref="GI:1245160"
                     /translation="MDVTIQHPWFKRALGPFYPSRLFDQFFGEGLFEYDLLPFLSSTI
                     SPYYRQSLFRTVLDSGISELMTHMWFVMHQPHAGNPKNNPGKVRSDRDKFVIFLDVKH
                     FSPEDLTVKVLEDFVEIHGKHNERQDDHGYISREFHRRYRLPSNVDQSALSCSLSADG
                     MLTFSGPKVQSGLDAGHSERAIPVSREEKPSSAPSS"
     misc_feature    348. .416
                     /note="alternatively spliced region; alpha A
                     (insert)-crystallin mRNA is produced by alternative
                     splicing from the alpha A-crystallin gene; insert peptide
                     sequence: LMTHMWFVMHQPHAGNPKNNPGK"
     3'UTR           750. .1271
ORIGIN

Alignment Scores:
Pred. No.:              9.13e-86        Length:         1271
Score:                  849.50          Matches:        164
Percent Similarity:     86.22%          Conservative:   5
Best Local Similarity:  83.67%          Mismatches:     4
Query Match:            92.74%          Indels:         23
DB:                     10              Gaps:           1

US-10-657-740-1 (1-173) x RNU47921 (1-1271)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        159 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 218

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 278

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 338
```

```
Qy          61 IleSerGlu---------------------------------------------------- 63
               |||||||||
Db         339 ATCTCTGAGCTCATGACCCATATGTGGTTTGTAATGCACCAACCACATGCTGGAAACCCC 398

Qy          64 ------------------ValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                                 ||||||||||||||||||||||||||||||||||||||||
Db         399 AAGAACAACCCCGGCAAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTG 458

Qy          78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
               ||||||||||||||||||||||||||||||||||||   :::|||||||||||||||||
Db         459 AAGCACTTCTCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCAT 518

Qy          98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         519 GGCAAACACAACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGC 578

Qy         118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         579 TACCGTCTGCCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGC 638

Qy         138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
               |||||||||||  ||||||||:::|||:::|||||||||||   |||:::||||||
Db         639 ATGCTGACCTTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGG 698

Qy         158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
               |||||||||||||||||||||||||||||:::|||||||||||||||
Db         699 GCCATTCCCGTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 746

RESULT 11
MUSCRYAAB
LOCUS       MUSCRYAAB               1025 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse lens alpha-A-crystallin mRNA, 3' end.
ACCESSION   J00376
VERSION     J00376.1  GI:192760
KEYWORDS    alpha-crystallin; crystallin.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   King,C.R., Shinohara,T. and Piatigorsky,J.
  TITLE     alpha A-crystallin messenger RNA of the mouse lens: more noncoding
            than coding sequences
  JOURNAL   Science 215 (4535), 985-987 (1982)
  MEDLINE   83119896
   PUBMED   7156978
COMMENT     Original source text: Mouse lens, cDNA to mRNA.
            The crystallins contribute 90% of the soluble lens protein and are
            highly conserved during evolution. There are four immunologically
            separate classes of crystallins called alpha- beta-, gamma-, and
            delta-crystallin. The mRNA of the alpha-A-crystallin from mouse,
            like that from cows and rats, is nearly three times the size
            required to code for the polypeptide. Nucleotides 1-490 of the
            mouse alpha-A-crystallin mRNA predict a protein sequence identical
            to amino acids 10-173 of rat alpha-A-crystallin.
FEATURES             Location/Qualifiers
     source          1. .1025
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     mRNA            <1. .1025
                     /product="CRY-alpha-A mRNA"
     CDS             <1. .492
                     /note="alpha-A-crystallin"
                     /codon_start=1
                     /protein_id="AAA37471.1"
                     /db_xref="GI:387134"
                     /translation="KRALGPFYPSRLFDQFFGEGLFEYDLLPFLSSTISPYYRQSLFR
                     TVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVLEDFVEIHGKHNERQDDHGYISR
                     EFHRRYRLPSNVDQSALSCSLSADGMLTFSGPKVQSGLDAGHSERAIPVSREEKPSSA
                     PSS"
ORIGIN      70 bp upstream of RsaI site.
```

```
Alignment Scores:
Pred. No.:                1.97e-81        Length:          1025
Score:                    810.00          Matches:         154
Percent Similarity:       97.55%          Conservative:    5
Best Local Similarity:    94.48%          Mismatches:      4
Query Match:              88.43%          Indels:          0
DB:                       10              Gaps:            0

US-10-657-740-1 (1-173) x MUSCRYAAB (1-1025)

Qy          11 LysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGly 30
               ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AAGCGTGCCCTGGGGCCCTTCTACCCCAGCCGACTGTTCGACCAGTTCTTCGGCGAGGGC 60

Qy          31 LeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGln 50
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CTTTTTGAGTACGACCTGCTGCCCTTCCTGTCTTCCACCATCAGCCCCTACTACCGCCAG 120

Qy          51 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 70
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCCCTCTTCCGCACTGTGCTGGACTCGGGCATCTCTGAGGTCCGTTCCGACCGGGACAAG 180

Qy          71 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 90
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TTTGTCATCTTCTTGGACGTGAAGCACTTCTCTCCTGAGGACCTCACCGTGAAAGTACTG 240

Qy          91 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 110
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAAGATTTTGTGGAGATTCACGGCAAACACAACGAGAGGCAGGATGACCATGGCTACATA 300

Qy         111 SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TCCCGTGAATTTCACCGTCGCTACCGTCTGCCTTCCAATGTGGACCAGTCCGCCCTCTCC 360

Qy         131 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 150
               |||||||||||||||||||||||||||||||||||  ||||||||:::|||:::||||||
Db         361 TGCTCCCTGTCTGCTGATGGCATGCTGACCTTCTCTGGCCCCAAAGTCCAGTCCGGTTTG 420

Qy         151 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 170
               ||||||  |||:::|||||||||||||||||||||||||||||||||:::||||||
Db         421 GATGCTGGCCACAGCGAGAGGGCCATTCCTGTGTCACGGGAGGAGAAACCCAGCTCTGCA 480

Qy         171 ProSerSer 173
               |||||||||
Db         481 CCCTCGTCC 489

RESULT 12
RCAACRYST
LOCUS       RCAACRYST               522 bp    mRNA    linear   VRT 18-SEP-1995
DEFINITION  R.catesbeiana mRNA for alpha-crystallin alpha A subunit.
ACCESSION   X85205
VERSION     X85205.1  GI:732915
KEYWORDS    aA-crystallin.
SOURCE      Rana catesbeiana (bullfrog)
  ORGANISM  Rana catesbeiana
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana;
            Aquarana.
REFERENCE   1
  AUTHORS   Lu,S.F., Pan,F.M. and Chiou,S.H.
  TITLE     Sequence analysis of frog alpha-crystallin cDNA and its deduced
            primary structure: comparison of alpha A subunit chains among
            different vertebrate species
  JOURNAL   Biochem. Biophys. Res. Commun. 210 (3), 974-981 (1995)
  MEDLINE   95283564
   PUBMED   7763271
REFERENCE   2  (bases 1 to 522)
  AUTHORS   Chiou,S.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-1995) S.H. Chiou, Institute of Biological
            Chemistry, Academia Sinica, Institute of Biochemical Sciences,
            National Taiwan University, P.O. Box 23-106, Taipei 10098, TAIWAN
COMMENT     Sequence overlapping with that under the acc#X00716.
```

```
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Rana catesbeiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:8400"
     CDS             1. .522
                     /codon_start=1
                     /product="alpha-crystallin alpha A subunit"
                     /protein_id="CAA59471.1"
                     /db_xref="GI:732916"
                     /db_xref="GOA:Q91311"
                     /db_xref="UniProt/Swiss-Prot:Q91311"
                     /translation="MDIAIQHPWFKRALGPFYPNRLFDQVFGEGMFDYDLFPFLSSTV
                     SPYYRHSLFRGFMDSGISEVRSDRDRFTINLDVKHFSPDDLTVKILDDFVEIHGKHSE
                     RQDDHGYISREFHRRYRLPSNLDQSSISCSLSADGILTFSGPKMMSGLDSSHSERPIP
                     VSREEKPTSAPSS"
ORIGIN

Alignment Scores:
Pred. No.:              5.48e-79        Length:         522
Score:                  785.00          Matches:        141
Percent Similarity:     92.49%          Conservative:   19
Best Local Similarity:  81.50%          Mismatches:     13
Query Match:            85.70%          Indels:         0
DB:                     5               Gaps:           0

US-10-657-740-1 (1-173) x RCAACRYST (1-522)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              |||||:::   |||||||||||||||||||||||    ||||||||||||||||||:::
Db          1 ATGGACATCGCCATCCAGCACCCCTGGTTCAAGCGTGCTCTGGGGCCCTTCTACCCCAAC 60

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||   ||||||||||:::|||:::|||||||   |||||||||
Db         61 CGCCTCTTCGATCAGGTCTTCGGAGAAGGAATGTTTGACTACGACCTCTTTCCTTTCCTG 120

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              |||||||||:::|||||||||||||    ||||||||||      :::|||||||||
Db        121 TCCTCCACCGTCAGTCCCTACTACAGGCACAGCCTCTTCCGAGGATTCATGGACTCCGGC 180

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              |||||||||||||||||||||||||:::|||   |||   ||||||||||||||||||
Db        181 ATTTCTGAGGTGCGTTCAGATCGCGATCGTTTCACCATTAACTTGGATGTGAAGCATTTC 240

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||:::|||||||||||||||:::   |||||||||||||||||||||||||||||
Db        241 TCACCCGATGATCTGACTGTGAAGATTCTGGATGACTTTGTGGAAATCCATGGAAAGCAC 300

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCGAGAGGCAGGATGACCATGGCTACATTTCCCGGGAGTTCCACCGCCGCTACCGCCTC 360

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              |||||||||:::|||||||||::::::||||||||||||||||||||||:::||||||
Db        361 CCCTCCAACCTGGACCAATCCTCTATCAGCTGCTCACTCTCTGCTGATGGCATTTTGACC 420

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::   :::||||||||||::::::|||:::||||||   ||||||
Db        421 TTCAGCGGACCAAAAATGATGTCAGGCCTGGACTCCAGCCACAGCGAGAGGCCTATCCCT 480

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCCAGAGAGGAAAAGCCCACCTCTGCCCCCTCCTCC 519

RESULT 13
D86299
LOCUS       D86299                  1361 bp    mRNA    linear   VRT 29-MAR-2003
DEFINITION  Cynops pyrrhogaster mRNA for newt alpha A-crystallin, complete cds.
ACCESSION   D86299
VERSION     D86299.1  GI:19168451
KEYWORDS    .
SOURCE      Cynops pyrrhogaster (Japanese firebelly newt)
  ORGANISM  Cynops pyrrhogaster
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
            Cynops.
REFERENCE   1
  AUTHORS   Mizuno,N., Agata,K., Sawada,K., Mochii,M. and Eguchi,G.
  TITLE     Expression of crystallin genes in embryonic and regenerating newt
            lenses
  JOURNAL   Dev. Growth Differ. 44 (3), 251-256 (2002)
  MEDLINE   22054970
   PUBMED   12060074
REFERENCE   2  (bases 1 to 1361)
  AUTHORS   Mizuno,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1996) Nobuhiko Mizuno, Osaka University,
            Institute for Molecular and Cellular Biology, Kondoh
            Differentiation Signaling Progect ERATO; 1-3 Yamadaoka, Suitashi,
            Osaka 565-0871, Japan (E-mail:nmizuno@imcb.osaka-u.ac.jp,
            Tel:06-6879-7964, Fax:06-6877-1738)
FEATURES             Location/Qualifiers
     source          1. .1361
                     /organism="Cynops pyrrhogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:8330"
     CDS             142. .663
                     /codon_start=1
                     /product="newt alpha A-crystallin"
                     /protein_id="BAB85811.1"
                     /db_xref="GI:19168452"
                     /translation="MDITIQHPWFKRALGPFYPGRLFDQFFGDGLFDYELFPFLSSTV
                     SPYYRQSMFRNYLDSGISEVRSGRDKFQIYLDVKHFSPEDLSVKILDDYVEIHGKHSD
                     RQDDHGYVSREFHRRYRLPASVDQSSITCSLSTDGMLTFSGTKMQSSLDSSHGERPIP
                     VSREEKPASAQSS"
ORIGIN

Alignment Scores:
Pred. No.:              1.58e-75        Length:         1361
Score:                  759.00          Matches:        135
Percent Similarity:     90.75%          Conservative:   22
Best Local Similarity:  78.03%          Mismatches:     16
Query Match:            82.86%          Indels:         0
DB:                     5               Gaps:           0

US-10-657-740-1 (1-173) x D86299 (1-1361)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||:::|||||||||||||||||||||||||    ||||||||||||||||
Db        142 ATGGACATCACCATCCAGCACCCATGGTTCAAACGGGCCCTGGGACCCTTTTACCCTGGT 201

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              |||||||||||||||||||||||:::|||||||||:::|||:::|||   |||||||||
Db        202 CGCCTCTTCGACCAGTTCTTCGGGGATGGCCTCTTCGACTACGAGCTCTTCCCCTTCTTA 261

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              |||||||||:::|||||||||||||||||||||:::||||||      ||||||||||||
Db        262 TCCTCCACCGTCAGTCCCTACTACCGCCAGTCCATGTTCCGCAACTACCTGGACTCTGGC 321

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||   ||||||||||   |||:::||||||||||||||||||
Db        322 ATATCCGAGGTGAGGTCTGGCCGCGACAAGTTCCAGATTTACCTGGACGTCAAGCACTTC 381

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||:::||||||:::   ||||||:::|||||||||||||||||||||
Db        382 TCACCGGAAGACCTAAGTGTCAAGATCCTCGACGATTACGTGGAGATCCATGGAAAGCAC 441

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ::::::||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        442 AGCGACCGGCAGGATGATCACGGGTATGTCTCGCGAGAGTTCCACCGCCGCTATCGCCTG 501

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              |||::::::||||||||||||:::::::::||||||||||   ||||||||||||||||
Db        502 CCCGCCAGCGTGGACCAGTCCTCCATCACATGCTCCCTGTCCACTGATGGCATGCTGACC 561

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
```

```
                ||| ||| |||:::|||::: ||||||::::::||| |||||| ||||||
Db        562 TTCTCGGGCACAAAGATGCAGTCTAGCTTGGACTCCAGCCACGGCGAGAGGCCCATTCCT 621

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||| |||||| ||||||
Db        622 GTTTCGCGAGAGGAGAAGCCCGCCTCTGCCCAATCCTCT 660

RESULT 14
AJ617725
LOCUS       AJ617725                447 bp    mRNA    linear   MAM 15-JUL-2004
DEFINITION  Elephas maximus partial mRNA for alphaA-crystallin (cryaA gene).
ACCESSION   AJ617725
VERSION     AJ617725.1  GI:50344346
KEYWORDS    alphaA-crystallin; cryaA gene.
SOURCE      Elephas maximus (Asiatic elephant)
  ORGANISM  Elephas maximus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
REFERENCE   1
  AUTHORS   van Rheede,T. and de Jong,W.W.
  TITLE     The alpha-crystallins of the platypus Ornithorhynchus anatinus:
            Origin of the alternatively spliced exon alphaAins and implications
            for mammalian phylogeny
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 447)
  AUTHORS   Franck,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-2003) Franck E., Biochemistry, Radboud University
            Nijmegen, 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Elephas maximus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9783"
     gene            1. .447
                     /gene="cryaA"
     CDS             <1. .>447
                     /gene="cryaA"
                     /codon_start=1
                     /product="alphaA-crystallin"
                     /protein_id="CAF02101.1"
                     /db_xref="GI:50344347"
                     /translation="RALGPFYPSRLFDQFFGEGLFEYDLLPFLSSTISPYYRQSLFRT
                     VLDSGISEVRSDRDQFLILLDVKHFSPEDLTVKVQDDFVEIHGKHNERQDDHGYISRE
                     FHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQSGMDASHSERAIP"
ORIGIN

Alignment Scores:
Pred. No.:              8.35e-76        Length:         447
Score:                  756.00          Matches:        141
Percent Similarity:     98.66%          Conservative:   6
Best Local Similarity:  94.63%          Mismatches:     2
Query Match:            82.53%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x AJ617725 (1-447)

Qy         12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCTCTGGGGCCCTTCTACCCTAGCCGGTTGTTTGATCAGTTTTTTGGTGAGGGCCTC 60

Qy         32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTGAGTACGACCTGCTGCCCTTCCTGTCCTCCACCATCAGCCCCTACTACCGCCAGTCC 120

Qy         52 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        121 CTCTTCCGCACCGTGCTCGACTCCGGCATCTCCGAGGTCCGGTCTGATCGGGACCAGTTC 180

Qy         72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91
              :::||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCATCCTCCTGGACGTGAAGCACTTCTCCCCTGAGGACCTGACTGTGAAGGTGCAGGAT 240

Qy         92 AspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACTTTGTGGAGATCCATGGCAAACACAATGAGAGGCAGGACGACCACGGCTACATCTCC 300

Qy        112 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTGAGTTCCACCGCCGCTACCGCCTGCCCTCCAACGTGGACCAGTCTGCACTCTCTTGC 360

Qy        132 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 151
              |||||||||||||||||||||||||||||||||||||||||||||||:::|||:::|||
Db        361 TCCCTGTCTGCGGACGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGTCTGGCATGGAT 420

Qy        152 AlaThrHisAlaGluArgAlaIlePro 160
              |||:::|||:::|||||||||||||||
Db        421 GCCAGCCACAGTGAGAGAGCCATCCCC 447

RESULT 15
D88185
LOCUS       D88185                  705 bp    mRNA    linear   VRT 14-APR-1999
DEFINITION  Xenopus laevis mRNA for alpha A crystallin, complete cds.
ACCESSION   D88185
VERSION     D88185.1  GI:4589827
KEYWORDS    alpha A crystallin.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (sites)
  AUTHORS   Mizuno,N., Mochii,M., Takahashi,T.C., Eguchi,G. and Okada,T.S.
  TITLE     Lens regeneration in Xenopus is not a mere repeat of lens
            development, with respect to crystallin gene expression
  JOURNAL   Differentiation 64 (3), 143-149 (1999)
  MEDLINE   99251119
   PUBMED   10234811
REFERENCE   2  (bases 1 to 705)
  AUTHORS   Mizuno,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-1996) Nobuhiko Mizuno, Biohistory Research Hall,
            Laboratory Div.; Murasaki chou 1-1, Takatuki, Osaka 569-11, Japan
            (E-mail:Nobuhiko.Mizuno@gate.brh.co.jp, Tel:0726-81-9751,
            Fax:0726-81-9757)
COMMENT     Sequence updated (12-Apr-1999).
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
     CDS             67. .582
                     /codon_start=1
                     /product="alpha A crystallin"
                     /protein_id="BAA76897.1"
                     /db_xref="GI:4589828"
                     /translation="MDITIQHPWFKRSLGPFYPNRLFDQVFGEGMFDFDLFPFMSSTI
                     SPYYKQNLSRGYLDSGISEVRSDRDRFVINLDVKHFSPEDLSVKVHDDFVEIHGKHNE
                     RQDDHGYISREFHRRYRLPSNMDQNSVSCTLSADGILTLFGPKLQSNMDSSHSDRTIP
                     VSKEEKSGSSS"
     polyA_site      705
                     /note="22 a nucleotides"
ORIGIN

Alignment Scores:
Pred. No.:              2.67e-73        Length:         705
Score:                  736.00          Matches:        131
Percent Similarity:     92.35%          Conservative:   26
Best Local Similarity:  77.06%          Mismatches:     13
Query Match:            80.35%          Indels:         0
DB:                     5               Gaps:           0

US-10-657-740-1 (1-173) x D88185 (1-705)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||:::|||||||||||||||||||||||||:::|||||||||||||||||:::
```

```
Db         67 ATGGATATCACCATTCAGCACCCCTGGTTCAAGCGCTCCCTGGGGCCCTTCTACCCCAAC 126

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40

Db        127 CGCCTGTTTGACCAGGTCTTTGGTGAGGGAATGTTTGACTTTGACCTGTTCCCCTTCATG 186

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60

Db        187 TCCTCCACCATCAGCCCCTACTACAAGCAGAACCTCTCCAGAGGATACCTGGACTCTGGG 246

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80

Db        247 ATCTCTGAGGTGCGTTCAGACCGTGACCGTTTTGTCATTAACCTGGATGTGAAGCATTTC 306

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100

Db        307 TCCCCTGAAGATCTGAGCGTCAAAGTCCATGATGACTTTGTAGAGATTCATGGGAAACAC 366

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120

Db        367 AACGAGCGACAGGATGATCACGGATACATCTCCCGAGAATTCCATCGCCGATATCGTCTT 426

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140

Db        427 CCTTCCAATATGGATCAGAACTCTGTGAGCTGCACTCTGTCTGCGGACGGGATCCTCACT 486

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160

Db        487 TTATTCGGTCCCAAACTGCAATCCAACATGGACTCCAGCCACAGCGATAGGACCATTCCT 546

Qy        161 ValSerArgGluGluLysProThrSerAla 170

Db        547 GTGTCCAAGGAGGAGAAATCAGGCTCATCC 576

RESULT 16
AJ617724
LOCUS       AJ617724                558 bp    mRNA    linear   MAM 15-JUL-2004
DEFINITION  Ornithorhynchus anatinus partial mRNA for alphaA-crystallin (cryaA
            gene).
ACCESSION   AJ617724
VERSION     AJ617724.1  GI:50344344
KEYWORDS    alphaA-crystallin; cryaA gene.
SOURCE      Ornithorhynchus anatinus (platypus)
  ORGANISM  Ornithorhynchus anatinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
REFERENCE   1
  AUTHORS   van Rheede,T. and de Jong,W.W.
  TITLE     The alpha-crystallins of the platypus Ornithorhynchus anatinus:
            Origin of the alternatively spliced exon alphaAins and implications
            for mammalian phylogeny
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 558)
  AUTHORS   Franck,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-2003) Franck E., Biochemistry, Radboud University
            Nijmegen, 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Ornithorhynchus anatinus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9258"
     gene            1. .558
                     /gene="cryaA"
     CDS             <1. .558
                     /gene="cryaA"
                     /codon_start=1
                     /product="alphaA-crystallin"
                     /protein_id="CAF02100.1"
                     /db_xref="GI:50344345"
                     /translation="RALGPLYPSRLFDQFFGEGLFEYDLLPFLSSTISPYYRQSLFRT
                     VLESGISELMAHVWFVMHKSHAGNPKNNPAKVRSDRDKFVIFLDVKHFSPEDLTVKVL
                     DEFVEIHGKHSERQDDHGYISREFHRRYRLPSNVDQASVSCSLSSDGMLTFSGSKVQS
                     NLDSSHSDRSIPISREEVPTSTPSS"
     exon            157. .225
                     /gene="cryaA"
                     /note="ins exon"
ORIGIN

Alignment Scores:
Pred. No.:              1.39e-72        Length:         558
Score:                  728.50          Matches:        139
Percent Similarity:     83.24%          Conservative:   15
Best Local Similarity:  75.14%          Mismatches:     8
Query Match:            79.53%          Indels:         23
DB:                     4               Gaps:           1

US-10-657-740-1 (1-173) x AJ617724 (1-558)

Qy         12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31

Db          1 CGTGCGTTGGGGCCTTTATACCCCAGCCGCCTGTTTGACCAGTTTTTTGGAGAGGGTCTC 60

Qy         32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51

Db         61 TTCGAGTACGACCTCCTGCCTTTCCTCTCCTCTACCATCAGTCCTTATTACCGGCAGTCT 120

Qy         52 LeuPheArgThrValLeuAspSerGlyIleSerGlu---------------------- 63

Db        121 CTGTTCCGCACCGTCCTGGAGTCGGGCATTTCCGAGCTCATGGCCCATGTATGGTTTGTA 180

Qy         64 ----------------------------------------------ValArgSerAspArg 68

Db        181 ATGCACAAATCACATGCTGGAAATCCCAAGAACAACCCTGCCAAGGTGCGATCTGACCGG 240

Qy         69 AspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLys 88

Db        241 GACAAGTTTGTCATCTTCCTGGATGTAAAGCATTTCTCTCCCGAAGATCTGACGGTGAAG 300

Qy         89 ValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGly 108

Db        301 GTGTTGGATGAGTTTGTGGAGATTCATGGCAAGCACAGTGAGAGACAGGATGACCACGGT 360

Qy        109 TyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAla 128

Db        361 TACATTTCCCGGGAATTCCATCGCCGGTACCGCCTGCCTTCCAACGTGGACCAGGCCTCA 420

Qy        129 LeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThr 148

Db        421 GTCTCCTGCTCCCTGTCCTCTGATGGCATGTTGACCTTCTCTGGCTCCAAGGTCCAATCC 480

Qy        149 GlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThr 168

Db        481 AACCTGGATTCCAGCCACAGTGACCGATCCATCCCTATCTCCCGGGAGGAAGTGCCCACC 540

Qy        169 SerAlaProSerSer 173

Db        541 TCAACTCCCTCTTCT 555

RESULT 17
AJ617726
LOCUS       AJ617726                477 bp    mRNA    linear   VRT 15-JUL-2004
DEFINITION  Sphenodon punctatus partial mRNA for alphaA-crystallin (cryaA
            gene).
ACCESSION   AJ617726
VERSION     AJ617726.1  GI:50344348
KEYWORDS    alphaA-crystallin; cryaA gene.
SOURCE      Sphenodon punctatus (tuatara)
  ORGANISM  Sphenodon punctatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
REFERENCE   1
  AUTHORS   van Rheede,T. and de Jong,W.W.
  TITLE     The alpha-crystallins of the platypus Ornithorhynchus anatinus:
            Origin of the alternatively spliced exon alphaAins and implications
            for mammalian phylogeny
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 477)
  AUTHORS   Franck,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-2003) Franck E., Biochemistry, Radboud University
            Nijmegen, 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Sphenodon punctatus"
                     /mol_type="mRNA"
                     /db_xref="taxon:8508"
     gene            1. .477
                     /gene="cryaA"
     CDS             <1. .477
                     /gene="cryaA"
                     /codon_start=1
                     /product="alphaA-crystallin"
                     /protein_id="CAF02102.1"
                     /db_xref="GI:50344349"
                     /translation="PLIPSRLFDQFFGEGLFEYDLLPLFSSTISPYYRQSFFRTVLES
                     GISEVRSDRDKFTIFLDVKHFSPEDLSVKIIDDFVEIHGKHNERQVDHGYISREFHRR
                     YRLPSNVDQSAITCSLSGDGMMTFSAPKVQSNMDPSHSERPIPVSREEKPTSAPSS"
ORIGIN

Alignment Scores:
Pred. No.:              3.79e-71        Length:         477
Score:                  715.00          Matches:        133
Percent Similarity:     91.14%          Conservative:   11
Best Local Similarity:  84.18%          Mismatches:     14
Query Match:            78.06%          Indels:         0
DB:                     5               Gaps:           0

US-10-657-740-1 (1-173) x AJ617726 (1-477)

Qy         16 ProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAsp 35
              |||      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCCTAATTCCAAGCCGTTTGTTCGACCAGTTTTTTGGAGAAGGTCTTTTTGAGTATGAT 60

Qy         36 LeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThr 55
              |||||||||       ||||||||||||||||||||||||||||||||||   |||||||||
Db         61 CTCCTGCCTTTGTTCTCTTCCACCATCAGCCCCTACTACAGGCAATCTTTCTTCCGCACC 120

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              |||||||:::|||||||||||||||||||||||||||||||||||||||   |||||||||
Db        121 GTCCTGGAATCAGGCATTTCAGAGGTGAGATCTGACCGGGACAAGTTTACAATCTTCCTG 180

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||||||||:::||||||:::   ||||||||||||||||
Db        181 GATGTAAAACACTTCTCTCCTGAAGATTTGAGTGTGAAGATTATTGATGACTTTGTGGAA 240

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        241 ATCCACGGCAAGCACAATGAGAGACAGGTAGACCATGGCTACATCTCCCGCGAATTCCAC 300

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              ||||||||||||||||||||||||||||||||||||||:::::|||||||||||||||||
Db        301 CGCAGGTACCGCCTGCCTTCCAATGTGGACCAATCCGCCATCACTTGCTCCCTGTCTGGT 360

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              |||||||||:::||||||      ||||||:::|||:::   :::|||   :::|||:::
Db        361 GACGGCATGATGACCTTCTCTGCCCCCAAGGTCCAGTCTAACATGGACCCCAGCCACAGC 420

Qy        156 GluArgAlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAGACCCATTCCTGTATCCCGTGAAGAGAAGCCCACCTCGGCGCCCTCCTCC 474

RESULT 18
AY035778
LOCUS       AY035778                730 bp    mRNA    linear   VRT 21-JUN-2002
DEFINITION  Danio rerio alpha A crystallin mRNA, complete cds.
ACCESSION   AY035778
VERSION     AY035778.1  GI:18266458
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Runkle,S., Hill,J., Kantorow,M., Horwitz,J. and Posner,M.
  TITLE     Sequence and spatial expression of zebrafish (Danio rerio)
            alphaA-crystallin
  JOURNAL   Mol. Vis. 8, 45-50 (2002)
  MEDLINE   21923047
   PUBMED   11925526
REFERENCE   2  (bases 1 to 730)
  AUTHORS   Runkle,S., Hill,J., Kantorow,M., Horwitz,J. and Posner,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2001) Biology, Ashland University, 401 College
            Avenue, Ashland, OH 44805, USA
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /tissue_type="lens"
     CDS             95. .616
                     /note="small heat shock protein"
                     /codon_start=1
                     /product="alpha A crystallin"
                     /protein_id="AAK61363.1"
                     /db_xref="GI:18266459"
                     /translation="MDIAIQHPWFRRTLGYPTRLFDQFFGEGLFDYDLFPFTTSTVSP
                     YYRHSLFRNILDSSNSGVSEVRSDREKFTVYLDVKHFSPDELSVKVTDDYVEIQGKHG
                     ERQDDHGYISREFHRRYRLPSNVDQSAITCTLSADGLLTLCGPKTSGIDAGRGDRTIP
                     VTREDKSNSGSSS"
ORIGIN

Alignment Scores:
Pred. No.:              1.51e-68        Length:         730
Score:                  694.00          Matches:        127
Percent Similarity:     85.80%          Conservative:   24
Best Local Similarity:  72.16%          Mismatches:     19
Query Match:            75.76%          Indels:         6
DB:                     5               Gaps:           3

US-10-657-740-1 (1-173) x AY035778 (1-730)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||:::   |||||||||||||||||:::|||||||||||      |||||||:::
Db         95 ATGGATATTGCGATCCAACACCCTTGGTTCAGACGCACACTGGGC------TACCCCACC 148

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||:::|||||||||   ||||||
Db        149 CGACTCTTTGATCAGTTCTTTGGAGAAGGCCTGTTCGATTATGACCTATTCCCCTTCACC 208

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp------ 58
              :::||||||:::|||||||||||||||   ||||||||||||   :::||||||
Db        209 ACTTCAACTGTCAGCCCTTACTATCGACACTCACTCTTCCGCAACATCCTGGACTCCTCC 268

Qy         59 ---SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                 ||||||:::|||||||||||||||||||:::|||||||   ::::::||||||||||
Db        269 AACTCGGGTGTCTCTGAGGTGAGGTCTGACAGAGAAAAATTTACAGTTTATCTGGATGTG 328

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||::::::|||:::||||||||||   ||||||:::||||||||||
Db        329 AAACATTTCTCTCCTGATGAGCTCAGTGTCAAGGTGACAGATGACTATGTGGAGATCCAG 388

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GGCAAGCATGGAGAAAGACAGGATGATCATGGCTACATCTCCCGTGAGTTCCATCGTCGC 448

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              ||||||||||||||||||||||||||||||||||||:::::|||:::||||||||||||||
Db        449 TACCGCCTGCCTTCCAATGTGGACCAGTCTGCCATCACCTGCACACTGTCTGCTGATGGC 508
```

```
Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              :::||||||   ||||||||||||      :::|||:::||||||         :::|||
Db        509 CTGCTCACTCTTTGTGGACCCAAG---ACAAGTGGCATAGATGCTGGCCGTGGAGATCGC 565

Qy        158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||:::||||||:::|||     |||      ||||||
Db        566 ACCATCCCTGTTACCCGCGAGGACAAGAGCAACTCAGGCTCTTCCTCC 613

RESULT 19
BC083177
LOCUS       BC083177               1430 bp    mRNA    linear   VRT 28-SEP-2004
DEFINITION  Danio rerio cDNA clone MGC:92036 IMAGE:7045051, complete cds.
ACCESSION   BC083177
VERSION     BC083177.1  GI:52789216
KEYWORDS    MGC.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1430)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1430)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Len Zon, Harvard
            cDNA Library Preparation: Open Biosystems
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 178 Row: p Column: 11
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23308654.
FEATURES    Location/Qualifiers
     source 1. .1430
            /organism="Danio rerio"
            /mol_type="mRNA"
            /db_xref="taxon:7955"
            /clone="MGC:92036 IMAGE:7045051"
            /tissue_type="Whole, adult zebrafish"
            /clone_lib="NIH_ZGC_10"
            /lab_host="DH10B"
            /note="Vector: pExpress1"
     CDS    106. .627
            /codon_start=1
            /product="Unknown (protein for MGC:92036)"
            /protein_id="AAH83177.1"
            /db_xref="GI:52789217"
            /translation="MDIAIQHPWFRRTLGYPTRLFDQFFGEGLFDYDLFPFTTSTVSP
            YYRHSLFRNILDSSNSGVSEVRSDREKFTVYLDVKHFSPDELSVKVTDDYVEIQGKHG
            ERQDDHGYISREFHRRYRLPSNVDQSAITCTLSADGLLTLCGPKTSGIDAGRGDRTIP
            VTREDKSNSGSSS"
ORIGIN

Alignment Scores:
Pred. No.:                3.55e-68        Length:          1430
Score:                    694.00          Matches:         127
Percent Similarity:       85.80%          Conservative:    24
Best Local Similarity:    72.16%          Mismatches:      19
Query Match:              75.76%          Indels:          6
DB:                       5               Gaps:            3

US-10-657-740-1 (1-173) x BC083177 (1-1430)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||:::   ||||||||||||||||||:::|||||||||||      |||||||:::
Db        106 ATGGATATTGCGATCCAACACCCTTGGTTCAGACGCACACTGGGC------TACCCCACC 159

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||:::|||||||||    ||||||
Db        160 CGACTCTTTGATCAGTTCTTTGGAGAAGGCCTGTTCGATTATGACCTATTCCCCTTCACC 219

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp------ 58
              :::||||||:::||||||||||||||||   ||||||||||||   :::||||||
Db        220 ACTTCAACTGTCAGCCCTTACTATCGACACTCACTCTTCCGCAACATCCTGGACTCCTCC 279

Qy         59 ---SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                 ||||||:::|||||||||||||||||||||||:::||||||   ||||||||||||||
Db        280 AACTCAGGTGTCTCTGAGGTGAGGTCTGACAGAGAAAAATTTACAGTTTACCTGGATGTG 339

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||:::::||:::||||||||||   ||||||:::||||||||||
Db        340 AAACATTTCTCTCCTGATGAGCTCAGTGTCAAGGTGACAGATGACTATGTGGAGATCCAG 399

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GGCAAGCATGGAGAAAGACAGGATGATCATGGCTACATCTCCCGTGAGTTCCATCGCCGC 459

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              |||||||||||||||||||||||||||||||||:::::|||:::||||||||||||||||
Db        460 TACCGCCTGCCTTCCAATGTGGACCAGTCTGCCATCACCTGCACACTGTCTGCTGATGGC 519

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              :::||||||   ||||||||||||      :::|||:::||||||         :::|||
Db        520 CTGCTCACTCTTTGTGGACCCAAG---ACAAGTGGCATAGATGCTGGCCGTGGAGATCGC 576

Qy        158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||:::||||||:::|||     |||      ||||||
Db        577 ACCATCCCTGTTACCCGCGAGGACAAGAGCAACTCGGGCTCTTCCTCC 624

RESULT 20
TSU31938
LOCUS       TSU31938               448 bp    mRNA    linear   VRT 25-SEP-1996
DEFINITION  Trachemys scripta elegans alpha A-crystallin mRNA, partial cds.
```

```
ACCESSION   U31938
VERSION     U31938.1  GI:1223846
KEYWORDS    .
SOURCE      Trachemys scripta elegans
  ORGANISM  Trachemys scripta elegans
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
REFERENCE   1  (bases 301 to 446)
  AUTHORS   Hedges,S.B., Simmons,M.D., van Dijk,M.A., Caspers,G.J., de
            Jong,W.W. and Sibley,C.G.
  TITLE     Phylogenetic relationships of the hoatzin, an enigmatic South
            American bird
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 92 (25), 11662-11665 (1995)
  MEDLINE   96102174
   PUBMED   8524824
REFERENCE   2  (bases 1 to 448)
  AUTHORS   Caspers,G.J.; Reinders,G.J., Leunissen,J.A., Wattel,J. and de
            Jong,W.W.
  TITLE     Protein sequences indicate that turtles branched off from the
            amniote tree after mammals
  JOURNAL   J. Mol. Evol. 42 (5), 580-586 (1996)
  MEDLINE   96259313
   PUBMED   8662010
REFERENCE   3  (bases 1 to 448)
  AUTHORS   de Jong,W.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-1995) Wilfried W. de Jong, Biochemistry,
            University of Nijmegen, P.O.Box 9101, 6500 HB Nijmegen, The
            Netherlands
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Trachemys scripta elegans"
                     /mol_type="mRNA"
                     /sub_species="elegans"
                     /db_xref="taxon:31138"
                     /tissue_type="lens"
     CDS             <1. .>448
                     /note="related to small heat shock proteins"
                     /codon_start=2
                     /product="alpha A-crystallin"
                     /protein_id="AAB08829.1"
                     /db_xref="GI:1223847"
                     /translation="RALGPLFPSRLFDQYLGEGLFDYDLLPFFSSTISPYYRHSLFRT
                     VLESGISEVRSDRDKFTILLDVKHFSPEDLSVKIMDDFVEIHGKHNERQDDHGYISRE
                     FHRRYRLPSNVDQSAITCSLSADGMLTFSGPKVQSNMDTSYSERPIP"
ORIGIN

Alignment Scores:
Pred. No.:              6.7e-67       Length:         448
Score:                  677.00        Matches:        123
Percent Similarity:     91.95%        Conservative:   14
Best Local Similarity:  82.55%        Mismatches:     12
Query Match:            73.91%        Indels:         0
DB:                     5             Gaps:           0

US-10-657-740-1 (1-173) x TSU31938 (1-448)

Qy          12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31
               |||   |||||||||   :::|||||||||||||||||||||:::   ||||||||||||
Db           2 CGAGCTCTTGGACCCTTATTTCCAAGCCGTTTGTTTGACCAGTATCTCGGAGAGGGTCTT 61

Qy          32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51
               |||:::|||||||||||||||||   |||||||||||||||||||||||||   |||
Db          62 TTCGATTATGATCTCCTGCCTTTTTTCTCTTCCACCATCAGCCCCTATTACAGGCACTCT 121

Qy          52 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 71
               ||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db         122 CTCTTCCGCACCGTTCTGGAATCAGGCATTTCAGAGGTGAGGTCTGACCGGGACAAGTTT 181

Qy          72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91
               |||   |||||||||||||||||||||||||||||||:::||||||:::   |||
Db         182 ACAATCCTCCTGGATGTAAAACACTTCTCTCCCGAAGATCTGAGTGTGAAGATTATGGAT 241

Qy          92 AspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 GACTTTGTGGAAATCCATGGCAAGCACAATGAGAGACAGGACGACCATGGCTACATTTCC 301

Qy         112 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 131
               ||||||||||||||||||||||||||||||||||||||||||||||||||:::::|||
Db         302 CGCGAATTCCACCGCAGATACCGCCTGCCTTCCAACGTGGACCAATCTGCCATCACCTGC 361

Qy         132 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 151
               |||||||||||||||||||||||||||||||   |||||||||:::|||:::   :::|||
Db         362 TCCCTGTCTGCTGATGGCATGCTGACTTTCTCTGGCCCAAAAGTCCAGTCCAACATGGAC 421

Qy         152 AlaThrHisAlaGluArgAlaIlePro 160
                   :::::::::||||||   ||||||
Db         422 ACCAGCTATAGCGAGAGACCCATTCCT 448

RESULT 21
AJ617727
LOCUS       AJ617727               447 bp    mRNA    linear   VRT 15-JUL-2004
DEFINITION  Lygodactylus picturatus partial mRNA for alphaA-crystallin (cryaA
            gene).
ACCESSION   AJ617727
VERSION     AJ617727.1  GI:50344350
KEYWORDS    alphaA-crystallin; cryaA gene.
SOURCE      Lygodactylus picturatus
  ORGANISM  Lygodactylus picturatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
            Lygodactylus.
REFERENCE   1
  AUTHORS   van Rheede,T. and de Jong,W.W.
  TITLE     The alpha-crystallins of the platypus Ornithorhynchus anatinus:
            Origin of the alternatively spliced exon alphaAins and implications
            for mammalian phylogeny
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 447)
  AUTHORS   Franck,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2003) Franck E., Biochemistry, Radboud University
            Nijmegen, 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Lygodactylus picturatus"
                     /mol_type="mRNA"
                     /db_xref="taxon:57064"
     gene            1. .447
                     /gene="cryaA"
     CDS             <1. .>447
                     /gene="cryaA"
                     /codon_start=1
                     /product="alphaA-crystallin"
                     /protein_id="CAF02103.1"
                     /db_xref="GI:50344351"
                     /translation="RAFGPLIPSRLFDQFFGEGLLDYDLLPLFSSTISPYYRQSLFRT
                     VLESGVSEVRSDRDKFTIFLDVKHFSPEDLSVKVIDDFVEIHGKHNERQDDHGYISRE
                     FHRRYRLPSNVDQSAISCSLSADGMLTFAAPKVQSNTDLSHSERPIP"
ORIGIN

Alignment Scores:
Pred. No.:              1.46e-66      Length:         447
Score:                  674.00        Matches:        125
Percent Similarity:     89.93%        Conservative:   9
Best Local Similarity:  83.89%        Mismatches:     15
Query Match:            73.58%        Indels:         0
DB:                     5             Gaps:           0

US-10-657-740-1 (1-173) x AJ617727 (1-447)

Qy          12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31
               |||      ||||||      ||||||||||||||||||||||||||||||||||||||
Db           1 CGGGCGTTTGGCCCCCTAATTCCAAGCCGTTTGTTCGACCAGTTTTTTGGAGAAGGTCTT 60

Qy          32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51
```

```
Db         61 CTTGATTATGATCTCCTGCCTTTGTTTTCCTCTACCATCAGCCCTTACTACAGGCAATCT 120

Qy         52 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 71

Db        121 CTCTTCCGCACTGTTCTTGAGTCAGGTGTTTCTGAGGTGAGATCTGACCGGGACAAGTTT 180

Qy         72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91

Db        181 ACAATCTTTTTGGATGTAAAACATTTCTCACCTGAAGATTTGAGCGTGAAGGTCATCGAT 240

Qy         92 AspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111

Db        241 GACTTTGTGGAAATCCATGGCAAACACAATGAGAGACAGGATGATCATGGGTACATCTCC 300

Qy        112 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 131

Db        301 CGTGAATTCCACCGCAGGTACCGCCTCCCTTCCAATGTGGACCAGTCGGCCATATCCTGC 360

Qy        132 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 151

Db        361 TCTCTGTCTGCTGATGGCATGCTGACCTTCGCTGCCCCTAAGGTCCAGTCCAACACTGAC 420

Qy        152 AlaThrHisAlaGluArgAlaIlePro 160

Db        421 CTTAGCCACAGTGAGAGACCCATTCCT 447

RESULT 22
EUDLAACRYS
LOCUS       EUDLAACRYS               448 bp    mRNA    linear   VRT 15-SEP-1994
DEFINITION  Eudromia elegans alpha A-crystallin mRNA, partial cds.
ACCESSION   L25850
VERSION     L25850.1  GI:538348
KEYWORDS    alpha A-crystallin.
SOURCE      Eudromia elegans (elegant crested-tinamou)
  ORGANISM  Eudromia elegans
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Palaeognathae; Tinamiformes; Tinamidae;
            Eudromia.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   Caspers,G.J., Wattel,J. and de Jong,W.W.
  TITLE     Alpha A-crystallin sequences group tinamou with ratites
  JOURNAL   Mol. Biol. Evol. 11 (4), 711-713 (1994)
  MEDLINE   94359400
   PUBMED   8078410
COMMENT     Original source text: Eudromia elegans eye lens cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Eudromia elegans"
                     /mol_type="mRNA"
                     /db_xref="taxon:8805"
                     /tissue_type="eye lens"
     CDS             <1. .>448
                     /codon_start=2
                     /product="alpha A-crystallin"
                     /protein_id="AAA49254.1"
                     /db_xref="GI:538349"
                     /translation="RALGPLIPSRLFDQFFGEGLLEYDLLPLFSSTISPYYRQSLFRS
                     VLESGISEVRSDREKFTIMLDVKHFSPEDLSVKIIDDFVEIHGKHSERQDDHGYISRE
                     FHRRYRLPSNVDQSAITCSLSSDGMLTFSGPKVQANMDPSHSERPIP"
ORIGIN

Alignment Scores:
Pred. No.:              2.53e-65       Length:        448
Score:                  663.00         Matches:       122
Percent Similarity:     90.60%         Conservative:  13
Best Local Similarity:  81.88%         Mismatches:    14
Query Match:            72.38%         Indels:        0
DB:                     5              Gaps:          0

US-10-657-740-1 (1-173) x EUDLAACRYS (1-448)

Qy         12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31

Db          2 CGGGCTCTGGGACCCCTGATTCCAAGCCGTTTGTTCGACCAGTTTTTCGGAGAGGGTCTC 61

Qy         32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51

Db         62 CTCGAGTACGATCTTCTGCCTCTGTTCTCTTCCACTATCAGCCCTTACTACAGGCAGTCT 121

Qy         52 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 71

Db        122 CTCTTCCGCAGCGTGCTGGAGTCAGGCATTTCAGAGGTGAGATCTGACCGGGAAAAATTT 181

Qy         72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91

Db        182 ACAATCATGCTGGATGTAAAACACTTCTCTCCTGAAGACCTAAGCGTGAAGATCATTGAT 241

Qy         92 AspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111

Db        242 GACTTTGTGGAAATCCATGGCAAGCACAGTGAAAGACAGGACGACCATGGCTACATCTCC 301

Qy        112 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 131

Db        302 CGTGAATTCCACCGCCGGTACCGCCTGCCCTCCAACGTGGACCAGTCTGCCATCACCTGC 361

Qy        132 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 151

Db        362 TCCCTCTCCAGTGACGGCATGCTGACCTTCTCCGGCCCCAAGGTCCAGGCCAACATGGAC 421

Qy        152 AlaThrHisAlaGluArgAlaIlePro 160

Db        422 CCCAGCCACAGCGAGAGACCCATCCCT 448

RESULT 23
APRAAC
LOCUS       APRAAC                   448 bp    mRNA    linear   VRT 09-SEP-2004
DEFINITION  A.platyrhynchos mRNA for alpha-A-crystallin.
ACCESSION   X96592
VERSION     X96592.1  GI:1945629
KEYWORDS    alpha-A-crystallin.
SOURCE      Anas platyrhynchos
  ORGANISM  Anas platyrhynchos
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
REFERENCE   1
  AUTHORS   Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
  TITLE     alpha-Crystallin sequences support a galliform/anseriform clade
  JOURNAL   Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
  MEDLINE   97271643
   PUBMED   9126559
REFERENCE   2  (bases 1 to 448)
  AUTHORS   Caspers,G.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
            Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Anas platyrhynchos"
                     /mol_type="mRNA"
                     /strain="Peking"
                     /db_xref="taxon:8839"
                     /tissue_type="lens"
     CDS             <1. .>448
                     /codon_start=2
                     /product="alpha-A-crystallin"
                     /protein_id="CAA65410.1"
                     /db_xref="GI:1945630"
                     /db_xref="GOA:O12984"
                     /db_xref="UniProt/Swiss-Prot:O12984"
                     /translation="RALGPLIPSRLFDQFFGEGLLEYDLLPLFSSTISPYYRQSLFRS
                     VLESGISEVRSDRDKFTIMLDVKHFSPEDLSVKIIDDFVEIHGKHSERQDDHGYISRE
                     FHRRYRLPANVDQSAITCSLSGDGMLTFSGPKVPSNMDPTHSERPIP"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              3.29e-65      Length:         448
Score:                  662.00        Matches:        122
Percent Similarity:     89.93%        Conservative:   12
Best Local Similarity:  81.88%        Mismatches:     15
Query Match:            72.27%        Indels:         0
DB:                     5             Gaps:           0

US-10-657-740-1 (1-173) x APRAAC (1-448)

Qy         12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31
              |||   |||||||||       |||||||||||||||||||||||||||||||||||||
Db          2 CGCGCTCTGGGACCCCTGATTCCCAGCCGTTTGTTCGACCAGTTTTTCGGAGAGGGTCTC 61

Qy         32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51
              ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db         62 CTGGAGTATGACCTCCTGCCTTTGTTCTCCTCCACTATCAGCCCCTACTACCGGCAGTCC 121

Qy         52 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 71
              ||||||||:::|||||:::|||||||||||||||||||||||||||||||||||||||||
Db        122 CTCTTCCGCAGCGTGCTGGAGTCGGGCATTTCAGAGGTGAGGTCTGACCGGGACAAGTTT 181

Qy         72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91
              |||   |||||||||||||||||||||||||||||:::||||||:::   |||
Db        182 ACGATCATGCTGGATGTAAAACACTTCTCTCCTGAAGATCTGAGCGTGAAGATTATCGAT 241

Qy         92 AspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        242 GACTTTGTGGAAATCCATGGCAAGCACAGTGAAAGACAGGATGACCACGGCTACATCTCC 301

Qy        112 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 131
              |||||||||||||||||||||||||||:::|||||||||||||||||:::::::|||
Db        302 CGTGAGTTTCACCGCCGGTACCGCCTGCCCGCCAACGTGGACCAGTCTGCCATCACCTGC 361

Qy        132 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 151
              |||||||||   |||||||||||||||||   |||||||||:::   :::   :::|||
Db        362 TCCCTCTCCGGCGACGGCATGCTGACCTTCTCAGGCCCCAAGGTCCCCTCCAACATGGAC 421

Qy        152 AlaThrHisAlaGluArgAlaIlePro 160
              |||||||:::||||||   ||||||
Db        422 CCCACCCACAGCGAGAGGCCCATCCCC 448

RESULT 24
CLRAAC
LOCUS       CLRAAC                   448 bp    mRNA    linear   VRT 09-SEP-2004
DEFINITION  C.livia mRNA for alpha-A-crystallin.
ACCESSION   X96593
VERSION     X96593.1  GI:1945730
KEYWORDS    alpha-A-crystallin.
SOURCE      Columba livia (domestic pigeon)
  ORGANISM  Columba livia
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
REFERENCE   1
  AUTHORS   Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
  TITLE     alpha-Crystallin sequences support a galliform/anseriform clade
  JOURNAL   Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
  MEDLINE   97271643
   PUBMED   9126559
REFERENCE   2  (bases 1 to 448)
  AUTHORS   Caspers,G.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
            Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Columba livia"
                     /mol_type="mRNA"
                     /db_xref="taxon:8932"
                     /tissue_type="lens"
     CDS             <1. .>448
                     /codon_start=2
                     /product="alpha-A-crystallin"
                     /protein_id="CAA65411.1"
                     /db_xref="GI:1945731"
                     /db_xref="GOA:O12988"
                     /db_xref="UniProt/Swiss-Prot:O12988"
                     /translation="RALGPLIPSRLFDQFFGEGLLEYDLLPWFSSTISPYYRQSLFRS
                     VLESGISEVRSDREKFTIMLDVKHFSPEDLSVKIIDDFVEIHGKHSERQDDHGYISRE
                     FHRRYRLPANVDQAAITCSLSNDGMLTFSGPKVPANMDASHGERPIP"
ORIGIN

Alignment Scores:
Pred. No.:              3.39e-64      Length:         448
Score:                  653.00        Matches:        120
Percent Similarity:     89.93%        Conservative:   14
Best Local Similarity:  80.54%        Mismatches:     15
Query Match:            71.29%        Indels:         0
DB:                     5             Gaps:           0

US-10-657-740-1 (1-173) x CLRAAC (1-448)

Qy         12 ArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeu 31
              |||   |||||||||       |||||||||||||||||||||||||||||||||||||
Db          2 CGTGCTCTGGGACCCCTCATTCCAAGCCGTTTGTTCGACCAGTTTTTTGGAGAGGGCCTC 61

Qy         32 PheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51
              ||||||||||||||||||:::   |||||||||||||||||||||||||||||||||||
Db         62 CTCGAGTACGATCTCCTGCCTTGGTTCTCTTCCACTATCAGCCCCTACTACAGGCAGTCC 121

Qy         52 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 71
              |||||||||:::|||||:::|||||||||||||||||||||||||||||:::||||||
Db        122 CTCTTCCGCAGTGTGCTGGAGTCAGGCATTTCAGAGGTGAGGTCTGACCGGGAAAAGTTT 181

Qy         72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91
              |||   |||||||||||||||||||||||||||||:::||||||:::   |||
Db        182 ACAATCATGCTGGATGTAAAACACTTCTCTCCCGAAGACCTGAGTGTGAAGATTATTGAC 241

Qy         92 AspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        242 GACTTTGTGGAAATCCATGGCAAGCACAGTGAAAGACAGGACGACCACGGTTACATCTCC 301

Qy        112 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 131
              |||||||||||||||||||||||||||:::|||||||||||:::|||:::::::|||
Db        302 CGCGAATTCCACCGCCGGTACCGCCTGCCCGCCAACGTGGACCAGGCTGCCATCACCTGC 361

Qy        132 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 151
              |||||||||   |||||||||||||||||   |||||||||:::         :::|||
Db        362 TCCCTGTCCAACGACGGCATGCTGACCTTCTCGGGCCCCAAGGTCCCCGCCAACATGGAC 421

Qy        152 AlaThrHisAlaGluArgAlaIlePro 160
              |||:::|||   ||||||   ||||||
Db        422 GCCAGCCACGGCGAGAGGCCCATCCCC 448

RESULT 25
AY007972
LOCUS       AY007972                 676 bp    mRNA    linear   VRT 23-OCT-2000
DEFINITION  Clarias fuscus alpha-A crystallin mRNA, complete cds.
ACCESSION   AY007972
VERSION     AY007972.1  GI:10946518
KEYWORDS    .
SOURCE      Clarias fuscus (whitespotted clarias)
  ORGANISM  Clarias fuscus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
            Clariidae; Clarias.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Chiou,S.-H. and Yu,C.-M.
  TITLE     Alpha crystallin of catfish eye lenes: cDNA and genomic analysis of
            alpha-A and alpha-B
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 676)
  AUTHORS   Chiou,S.-H. and Yu,C.-M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2000) Institute of Biological Chemistry, Academia
```

```
          Sinica, P.O. Box 23-106, Taipei, Taiwan 10098, Republic of China
FEATURES             Location/Qualifiers
     source          1. .676
                     /organism="Clarias fuscus"
                     /mol_type="mRNA"
                     /db_xref="taxon:33541"
     CDS             136. .657
                     /codon_start=1
                     /product="alpha-A crystallin"
                     /protein_id="AAG23866.1"
                     /db_xref="GI:10946519"
                     /translation="MDIAIQYPWFRRTLSNPSRLFDQFFGEGLLDHDLLPFTASTISP
                     SYRHSLFRSFLDSSNSGISEVRSDRDRFMVYLDVKHFSPEELRVKVADDYVEIHGKHG
                     ERQDDHGYISREFHRRYRLPSNVDQAAITCTLSADGLLSFCGPKTGGSKYGRGDRTIP
                     VTRDDKPNPAASS"
ORIGIN

Alignment Scores:
Pred. No.:              5.72e-64        Length:         676
Score:                  653.00          Matches:        124
Percent Similarity:     83.52%          Conservative:   23
Best Local Similarity:  70.45%          Mismatches:     23
Query Match:            71.29%          Indels:         6
DB:                     5               Gaps:           3

US-10-657-740-1 (1-173) x AY007972 (1-676)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||:::   ||||||:::|||||||||:::|||||||||            ||||||
Db        136 ATGGATATTGCCATCCAATACCCGTGGTTCAGACGTACCCTGAGC------AATCCCTCT 189

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||   ::::::||||||||||||||||
Db        190 CGACTCTTTGACCAATTCTTTGGAGAAGGTCTGCTTGATCATGACCTCCTCCCCTTTACT 249

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp------ 58
              :::|||||||||||||||   ||||||   ||||||||||||:::   ||||||
Db        250 GCCTCCACCATCAGCCCTTCCTACCGACACTCTCTCTTCCGCAGTTTTTTGGACTCCTCC 309

Qy         59 ---SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                 ||||||||||||||||||||||||||||||:::|||:::::::::|||||||||
Db        310 AACTCTGGCATTTCCGAGGTGAGGTCTGACAGAGACAGATTCATGGTTTACCTGGATGTG 369

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||||:::|||   |||||||||   ||||||:::||||||||||||
Db        370 AAGCACTTCTCCCCAGAGGAACTCAGAGTTAAGGTAGCAGATGACTATGTGGAGATTCAT 429

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GGCAAGCATGGAGAAAGACAGGACGACCATGGCTACATCTCACGGGAGTTCCACCGCCGC 489

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              ||||||||||||||||||||||||||:::|||:::::::|||:::|||||||||||||||
Db        490 TACCGTCTACCCTCCAACGTCGATCAGGCAGCCATCACTTGTACGCTGTCAGCTGATGGC 549

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              :::|||:::|||||||||||||||   |||          :::|||
Db        550 CTGCTCAGCTTCTGCGGCCCCAAA---ACAGGCGGATCAAAATACGGGCGTGGAGATCGC 606

Qy        158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||:::|||::::::||||||      |||   ||||||
Db        607 ACCATTCCTGTCACCCGCGATGACAAACCCAACCCTGCTGCTTCCTCT 654

RESULT 26
RTCRYA
LOCUS       RTCRYA                    765 bp    mRNA    linear   VRT 06-JUL-1989
DEFINITION  Frog mRNA fragment for alpha-A2-crystallin.
ACCESSION   X00716 X00058
VERSION     X00716.1  GI:64292
KEYWORDS    alpha-crystallin; crystallin.
SOURCE      Rana temporaria (common frog)
  ORGANISM  Rana temporaria
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana;
            Rana.
REFERENCE   1
  AUTHORS   Tomarev,S.I., Zinovieva,R.D., Dolgilevich,S.M., Krayev,A.S.,
            Skryabin,K.G. and Gause,G.G. Jr.
  TITLE     The absence of the long 3'-non-translated region in mRNA coding for
            eye lens alpha A2-crystallin of the frog (Rana temporaria)
  JOURNAL   FEBS Lett. 162 (1), 47-51 (1983)
  MEDLINE   84005173
   PUBMED   6604666
REFERENCE   2  (bases 1 to 765)
  AUTHORS   Tomarev,S.I., Zinov'eva,R.D., Kraev,A.S., Skryabin,K.G. and
            Gauze,G.G.
  TITLE     Primary structure of cloned cDNA coding alpha-A2 crystallin of the
            eye lens of the frog Rana temporaria
  JOURNAL   Dokl. Biochem. 271, 277-280 (1984)
COMMENT     Data kindly reviewed (05-11-1985) by S.I. Tomarev
            Sequence 1 to 185 is complementary to sequence 553 to 369 and
            probably a cloning artefact; amino terminal amino acids are
            missing.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Rana temporaria"
                     /mol_type="mRNA"
                     /db_xref="taxon:8407"
     misc_feature    1. .185
                     /note="artefactual sequence"
     CDS             186. .632
                     /note="unnamed protein product; Protein sequence is in
                     conflict with the conceptual translation; alpha-A2
                     crystalline (aa 25 to 173)"
                     /codon_start=1
                     /protein_id="CAA25308.1"
                     /db_xref="GI:1334774"
                     /db_xref="GOA:P02508"
                     /db_xref="UniProt/Swiss-Prot:P02508"
                     /translation="QVFGEGMFDYDLFPFLTSTVSPHYRHGLLRGFMDSGISEVRSDR
                     DRFTINLDVKHFSPDDLTVKILDDFVEIHGKHSERQDDHGYISREFHRRYRLPSNLDQ
                     SSISCSLSADGILTFSGPKMMSNLVSSHSERPIPVSREEKPTSAPSS"
     misc_feature    744. .749
                     /note="polyadenylation signal"
     polyA_site      765
                     /note="polyadenylation site"
ORIGIN

Alignment Scores:
Pred. No.:              2.02e-61        Length:         765
Score:                  631.00          Matches:        116
Percent Similarity:     88.82%          Conservative:   19
Best Local Similarity:  76.32%          Mismatches:     17
Query Match:            68.89%          Indels:         0
DB:                     5               Gaps:           0

US-10-657-740-1 (1-173) x RTCRYA (1-765)

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSer 41
              |||   |||   ||||||||||||:::|||:::|||||||||   |||||||||:::
Db        177 CTTCACAGTCAGGTCTTCGGAGAAGGGATGTTTGACTACGACCTCTTCCCTTTCCTGACC 236

Qy         42 SerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIle 61
              ||||||:::||||||:::||||||      |||   |||      :::||||||||||||
Db        237 TCCACCGTCAGCCCCCACTACCGGCACGGCCTCCTCCGAGGATTCATGGACTCCGGCATC 296

Qy         62 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 81
              ||||||||||||||||||||||||||:::|||   |||   ||||||||||||||||||
Db        297 TCTGAGGTGCGTTCAGATCGCGATCGTTTCACCATTAACTTGGACGTGAAGCATTTCTCC 356

Qy         82 ProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsn 101
              |||:::|||||||||||||||:::   |||||||||||||||||||||||||||||:::
Db        357 CCCGATGATCTGACTGTGAAGATTCTGGATGACTTTGTGGAAATCCATGGAAAGCACAGC 416

Qy        102 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        417 GAGAGGCAGGATGACCATGGCTACATATCCCGGGAGTTCCACCGCCGCTACCGCCTCCCC 476

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
              ||||||:::|||||||||:::::|||||||||||||||||||||||:::|||||||||
Db        477 TCCAACCTGGACCAATCCTCTATCAGCTGCTCACTCTCTGCCGATGGCATTCTGACCTTC 536

Qy        142 CysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 161
              |||||||||:::   :::   |||   ::::::|||:::||||||   |||||||||
Db        537 AGCGGACCAAAAATGATGTCAAACCTGGTCTCCAGCCACAGCGAGAGGCCCATCCCTGTG 596

Qy        162 SerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||
Db        597 TCCAGAGAGGAAAAGCCCACCTCTGCCCCCTCCTCC 632

RESULT 27
RNACRY
LOCUS       RNACRY                   885 bp    mRNA    linear   ROD 08-JAN-1999
DEFINITION  Partial sequence of messenger RNA for rat alpha-a2-crystallin.
ACCESSION   V01219 J00715
VERSION     V01219.1  GI:55598
KEYWORDS    complementary DNA; crystallin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Moormann,R.J., van der Velden,H.M., Dodemont,H.J., Andreoli,P.M.,
            Bloemendal,H. and Schoenmakers,J.G.
  TITLE     An unusually long non-coding region in rat lens alpha-crystallin
            messenger RNA
  JOURNAL   Nucleic Acids Res. 9 (19), 4813-4822 (1981)
  MEDLINE   82081811
   PUBMED   6171772
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     mRNA            <1. .885
     CDS             <1. .368
                     /codon_start=3
                     /product="crystallin"
                     /protein_id="CAA24530.1"
                     /db_xref="GI:809074"
                     /translation="FRTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVLEDFVEI
                     HGKHNERQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFSGPKVQSGLDAGH
                     SERAIPVSREEKPSSAPSS"
     polyA_site      885
ORIGIN

Alignment Scores:
Pred. No.:              7.83e-57        Length:         885
Score:                  591.00          Matches:        113
Percent Similarity:     97.52%          Conservative:   5
Best Local Similarity:  93.39%          Mismatches:     3
Query Match:            64.52%          Indels:         0
DB:                     10              Gaps:           0

US-10-657-740-1 (1-173) x RNACRY (1-885)

Qy         53 PheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheVal 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 TTCCGCACAGTGTTGGACTCCGGCATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTC 62

Qy         73 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||   :::|||
Db         63 ATCTTCTTGGATGTGAAGCACTTCTCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGAT 122

Qy         93 PheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArg 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTCGTGGAGATCCATGGCAAACACAACGAGAGGCAGGATGACCATGGCTACATTTCCCGT 182

Qy        113 GluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAATTTCACCGTCGCTACCGTCTGCCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCC 242

Qy        133 LeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAla 152
              |||||||||||||||||||||||||||   |||||||||:::|||:::||||||||||||
Db        243 TTGTCTGCGGATGGCATGCTGACCTTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCT 302

Qy        153 ThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAlaProSer 172
              |||:::|||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db        303 GGCCACAGCGAGAGGGCCATTCCCGTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCG 362

Qy        173 Ser 173
              |||
Db        363 TCC 365

RESULT 28
OLJ000940
LOCUS       OLJ000940                436 bp    mRNA    linear   VRT 02-MAY-1998
DEFINITION  Oryzias latipes, alpha-A-crystallin gene.
ACCESSION   AJ000940
VERSION     AJ000940.1  GI:3115327
KEYWORDS    alpha-A-crystallin gene.
SOURCE      Oryzias latipes (Japanese medaka)
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1
  AUTHORS   Loosli,F., Koster,R.W., Carl,M., Krone,A. and Wittbrodt,J.
  TITLE     Six3, a medaka homologue of the Drosophila homeobox gene sine
            oculis is expressed in the anterior embryonic shield and the
            developing eye
  JOURNAL   Mech. Dev.
REFERENCE   2  (bases 1 to 436)
  AUTHORS   Loosli,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-1998) Loosli F., Institute for Human Genetics,
            c/o MPI of Biophys. Chem., University of Gottingen, Am Fassberg,
            Gottingen, 37077, GERMANY
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Oryzias latipes"
                     /mol_type="mRNA"
                     /strain="Carolina Biological"
                     /db_xref="taxon:8090"
     gene            1. .436
                     /gene="alpha-A-crystallin"
     CDS             <1. .>436
                     /gene="alpha-A-crystallin"
                     /codon_start=1
                     /protein_id="CAA04397.1"
                     /db_xref="GI:3115328"
                     /db_xref="GOA:O73919"
                     /db_xref="UniProt/Swiss-Prot:O73919"
                     /translation="RLFDQFFGEGMFDHDLLPFTSPTISPFYRQSLFRNFLDSSNSGI
                     SEVRSDRDKFTVHWDVKHFSPDELSVKVIDDFVEIQGKHGERQDDHGYISREFHRRYR
                     LPSTVDQSAITCSLSADGLLTLSGPNPAGGPNGRSDRSIPVCR"
ORIGIN

Alignment Scores:
Pred. No.:              6.31e-52        Length:         436
Score:                  544.00          Matches:        106
Percent Similarity:     82.88%          Conservative:   15
Best Local Similarity:  72.60%          Mismatches:     21
Query Match:            59.39%          Indels:         4
DB:                     5               Gaps:           2

US-10-657-740-1 (1-173) x OLJ000940 (1-436)

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||:::|||::::::|||||||||||||||||
```

```
Db          1 CGGCTGTTCGACCAGTTTTTCGGGGAGGGGATGTTCGATCACGACCTCCTGCCCTTCACC 60

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp------ 58
              |||   |||||||||||:::|||||||||||||||||||||      ||||||
Db         61 TCCCCCACCATCAGTCCGTTCTACAGACAGTCGCTGTTTCGTAACTTCCTGGACTCCTCA 120

Qy         59 ---SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                 ||||||||||||||||||||||||||||||||||||||||||   :::      ||||||
Db        121 AACTCCGGCATTTCCGAGGTGAGGTCTGACAGGGACAAGTTCACGGTTCACTGGGATGTC 180

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||:::::|||:::|||||||||   ||||||||||||||||||
Db        181 AAGCACTTCTCCCCCGATGAGCTCAGCGTGAAGGTGATCGATGACTTTGTGGAGATCCAG 240

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCAAGCATGGAGAAAGACAGGATGACCATGGCTACATCTCCCGGGAGTTTCACCGCCGC 300

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              ||||||||||||||   ||||||||||||||::::::|||||||||||||||||||||||
Db        301 TACCGCCTCCCCTCCACTGTGGACCAATCGGCCATCACCTGCTCCCTGTCTGCCGATGGA 360

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              :::||||||      ||||||              |||   :::          ::::::|||
Db        361 CTGCTGACCCTGTCCGGGCCAAACCCAGCGGGGGGGCCGAATGGA---CGCAGCGACCGC 417

Qy        158 AlaIleProValSerArg 163
              :::|||||||||   |||
Db        418 AGCATCCCCGTCTGCCGC 435

RESULT 29
AF029793
LOCUS       AF029793                632 bp    mRNA    linear   MAM 30-JUN-1999
DEFINITION  Bos taurus alpha B-crystallin (CRYAB) mRNA, complete cds.
ACCESSION   AF029793
VERSION     AF029793.2  GI:5296002
KEYWORDS    .
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1  (bases 1 to 632)
  AUTHORS   Kelley,P.B., Abraham,E.C., Zhao,H.R., Shroff,N.P., Cherian,M. and
            Thomas,J.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-1997) Biochemistry and Molecular Biology, Medical
            College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
REFERENCE   2  (bases 1 to 632)
  AUTHORS   Kelley,P.B., Abraham,E.C., Zhao,H.R., Shroff,N.P., Cherian,M. and
            Thomas,J.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-1999) Biochemistry and Molecular Biology, Medical
            College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
  REMARK    Sequence update by submitter
COMMENT     On Jun 30, 1999 this sequence version replaced gi:2760900.
FEATURES             Location/Qualifiers
     source          1. .632
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
     gene            1. .632
                     /gene="CRYAB"
     CDS             1. .528
                     /gene="CRYAB"
                     /codon_start=1
                     /product="alpha B-crystallin"
                     /protein_id="AAB95323.2"
                     /db_xref="GI:5296003"
                     /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPASTSL
                     SPFYLRPPSFLRAPSWIDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                     KHEERQDEHGFISREFHRKYRIPADVDPLAITSSLSSDGVLTVNGPRKQASGPERTIP
                     ITREEKPAVTAAPKK"

ORIGIN

Alignment Scores:
Pred. No.:              1.19e-46      Length:          632
Score:                  499.00        Matches:         99
Percent Similarity:     75.28%        Conservative:    35
Best Local Similarity:  55.62%        Mismatches:      32
Query Match:            54.48%        Indels:          12
DB:                     4             Gaps:            6

US-10-657-740-1 (1-173) x AF029793 (1-632)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          |||||||:::   |||
Db          1 ATGGATATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTCCCTTTCCACTCTCCC 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         61 AGCCGCCTCTTTGACCAGTTTTTTGGCGAGCACCTGTTGGAGTCTGATCTCTTCCCA--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||       :::
Db        118 GCTTCTACTTCCCTGAGCCCCTTCTACCTTCGGCCGCCCTCATTTCTGCGGGCACCCAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        178 TGGATTGACACTGGCCTCTCAGAGATGCGTCTGGAGAAGGACAGATTCTCTGTCAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   ||||||||      |||   :::|||
Db        238 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAGGTCAAGGTGCTGGGAGATGTGATTGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::|||||:::||||||||||||||||||
Db        298 GTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGGGAGTTCCAC 357

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::|||||      |||::::::   |||||||||:::
Db        358 AGGAAATACCGGATCCCAGCTGACGTGGACCCTCTCGCCATTACTTCATCCCTGTCCTCT 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                |||:::
Db        418 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCCGGCCCT 465

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
Db        466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCGGCTGTCACTGCAGCCCCC 519

RESULT 30
OCCRYAB
LOCUS       OCCRYAB                 548 bp    mRNA    linear   MAM 23-OCT-1996
DEFINITION  O.cuniculus mRNA for alpha-B-crystallin.
ACCESSION   X95383
VERSION     X95383.1  GI:1177578
KEYWORDS    alpha-B-crystallin; cryaB gene.
SOURCE      Oryctolagus cuniculus (rabbit)
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE   1
  AUTHORS   Krausz,E., Augusteyn,R.C., Quinlan,R.A., Reddan,J.R., Russell,P.,
            Sax,C.M. and Graw,J.
  TITLE     Expression of Crystallins, Pax6, Filensin, CP49, MIP, and MP20 in
            lens-derived cell lines
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 37 (10), 2120-2128 (1996)
  MEDLINE   96409169
   PUBMED   8814151
REFERENCE   2  (bases 1 to 548)
  AUTHORS   Graw,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-1996) J. Graw, Institute of Mammalian Genetics,
            GSF-Research Center Neuherberg, Ingolstaedter Landstrasse 1,
            Oberschleissheim, D-85764, FRG
```

```
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Oryctolagus cuniculus"
                     /mol_type="mRNA"
                     /sub_species="New Zeeland white"
                     /db_xref="taxon:9986"
                     /tissue_type="lens"
     gene            1. .548
                     /gene="cryab"
     CDS             21. .548
                     /gene="cryab"
                     /codon_start=1
                     /product="alpha-B-crystallin"
                     /protein_id="CAA64669.1"
                     /db_xref="GI:1177579"
                     /db_xref="GOA:P41316"
                     /db_xref="UniProt/Swiss-Prot:P41316"
                     /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSL
                     SPFYLRPPSFLRAPSWIDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                     KHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQAPGPERTIP
                     ITREEKPAVTAAPKK"
ORIGIN

Alignment Scores:
Pred. No.:                4.72e-46       Length:          548
Score:                    493.00         Matches:         98
Percent Similarity:       74.16%         Conservative:    34
Best Local Similarity:    55.06%         Mismatches:      34
Query Match:              53.82%         Indels:          12
DB:                       4              Gaps:            6

US-10-657-740-1 (1-173) x OCCRYAB (1-548)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||  :::|||         ||||||:::   |||
Db         21 ATGGACATCGCTATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTTCACTCGCCC 80

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||  .|||
Db         81 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTCTTCCCA--- 137

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||       :::
Db        138 ACTTCTACTTCCCTGAGCCCCTTCTATCTTCGGCCACCCTCATTCCTGCGGGCACCCAGC 197

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        198 TGGATTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 257

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        258 GATGTGAAGCACTTCTCCCCAGAGGAGCTCAAGGTCAAAGTGTTGGGTGATGTGATTGAG 317

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        318 GTGCACGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 377

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        378 AGGAAATACCGGATCCCAGCTGATGTGGACCCTCTCACCATTACTTCATCCCTGTCATCT 437

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||              |||
Db        438 GATGGGGTCCTCACTGTGAATGGACCAAGGAAGCAA-----------GCCCCTGGCCCA 485

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        486 GAGCGCACCATTCCCATAACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCC 539

RESULT 31
BT006770
LOCUS       BT006770                 528 bp    mRNA    linear   PRI 13-MAY-2003
DEFINITION  Homo sapiens crystallin, alpha B mRNA, complete cds.
ACCESSION   BT006770
VERSION     BT006770.1  GI:30582378
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Cloning of human full-length CDSs in BD Creator(TM) System Donor
            vector
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 528)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
            Circle, Palo Alto, CA 94303, USA
COMMENT     This CDS clone is a part of a collection of human full length
            expression clones generated by BD Biosciences Clontech and the
            Harvard Institute of Proteomics. Each CDS has been cloned in two
            forms: with and without stop-codon (to allow fusion with C-terminal
            tag). The CDS has been directionally cloned using BD In-Fusion(TM)
            cloning system between the SalI and HindIII sites of the pDNR-DUAL
            vector. Additional sequences in the clone: 'ACC' after SalI site
            and before 'ATG' to provide Kozak consensus sequence; 'GG' after
            last codon and before HindIII site to maintain reading frame.
            Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GH00548X1.0"
                     /clone_lib="BD Creator(TM) CDS Library derived from MGC
                     collection"
                     /lab_host="DH5alpha T1 resistant"
                     /note="Vector: pDNR-Dual"
     CDS             1. .528
                     /codon_start=1
                     /product="crystallin, alpha B"
                     /protein_id="AAP35416.1"
                     /db_xref="GI:30582379"
                     /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSL
                     SPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                     KHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVSGPERTIP
                     ITREEKPAVTAAPKK"
ORIGIN

Alignment Scores:
Pred. No.:                1.27e-45       Length:          528
Score:                    489.00         Matches:         97
Percent Similarity:       73.60%         Conservative:    34
Best Local Similarity:    54.49%         Mismatches:      35
Query Match:              53.38%         Indels:          12
DB:                       9              Gaps:            6

US-10-657-740-1 (1-173) x BT006770 (1-528)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          1 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||       :::
Db        118 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 177
```

```
Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         178 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAAAAGGACAGGTTCTCTGTCAACCTG 237

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         238 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 297

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db         298 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 357

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db         358 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 417

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||      ||||||:::   |||     .          :::
Db         418 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 465

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db         466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 519

RESULT 32
BT007909
LOCUS       BT007909                 528 bp    mRNA    linear   SYN 13-MAY-2003
DEFINITION  Synthetic construct Homo sapiens crystallin, alpha B mRNA, partial
            cds.
ACCESSION   BT007909
VERSION     BT007909.1  GI:30584656
KEYWORDS    FLI_CDNA.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Cloning of human full-length CDSs in BD Creator(TM) System Donor
            vector
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 528)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
            Circle, Palo Alto, CA 94303, USA
COMMENT     This CDS clone is a part of a collection of human full length
            expression clones generated by BD Biosciences Clontech and the
            Harvard Institute of Proteomics. Each CDS has been cloned in two
            forms: with and without stop-codon (to allow fusion with C-terminal
            tag). The CDS has been  directionally cloned using BD In-Fusion(TM)
            cloning system between the SalI  and HindIII sites of the pDNR-DUAL
            vector. Additional sequences in the clone:  'ACC' after SalI site
            and before 'ATG' to provide Kozak consensus sequence; 'GG' after
            last codon and before HindIII site to maintain reading frame.
            Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /clone="GH00548L1.0"
                     /clone_lib="BD Creator(TM) CDS Library derived from MGC
                     collection"
                     /lab_host="DH5alpha T1 resistant"
                     /note="Vector: pDNR-Dual"
     CDS             1. .>528
                     /note="Mutations: 527:Stop->Leu"
                     /codon_start=1
                     /transl_table=11
                     /product="Homo sapiens crystallin, alpha B"
                     /protein_id="AAP36581.1"
                     /db_xref="GI:30584657"
                     /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSL
                     SPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                     KHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVSGPERTIP
                     ITREEKPAVTAAPKKL"
ORIGIN

Alignment Scores:
Pred. No.:              1.27e-45      Length:         528
Score:                  489.00        Matches:        97
Percent Similarity:     73.60%        Conservative:   34
Best Local Similarity:  54.49%        Mismatches:     35
Query Match:            53.38%        Indels:         12
DB:                     12            Gaps:           6

US-10-657-740-1 (1-173) x BT007909 (1-528)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   ||||||||   :::|||         ||||||:::   |||
Db           1 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 60

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               |||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db          61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 117

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  |||:::::::::|||||:::|||   :::      |||      |||      :::
Db         118 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 177

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         178 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAAAAGGACAGGTTCTCTGTCAACCTG 237

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         238 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 297

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db         298 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 357

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db         358 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 417

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||      ||||||:::   |||     .          :::
Db         418 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 465

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db         466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 519

RESULT 33
AX937702
LOCUS       AX937702                 537 bp    DNA     linear   PAT 06-JAN-2004
DEFINITION  Sequence 12 from Patent WO03091266.
ACCESSION   AX937702
VERSION     AX937702.1  GI:40713713
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Quinlan,R.U.
  TITLE     Protein stabilisation
  JOURNAL   Patent: WO 03091266-A 12 06-NOV-2003;
            The University Court of the University of Dundee (GB)
FEATURES             Location/Qualifiers
```

```
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.3e-45        Length:         537
Score:                  489.00         Matches:        97
Percent Similarity:     73.60%         Conservative:   34
Best Local Similarity:  54.49%         Mismatches:     35
Query Match:            53.38%         Indels:         12
DB:                     6              Gaps:           6

US-10-657-740-1 (1-173) x AX937702 (1-537)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          9 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 68

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||   |||   |||   |||||||   |||
Db         69 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 125

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||      :::
Db        126 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 185

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        186 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 245

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        246 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 305

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::|||||||||||||||||||
Db        306 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 365

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::|||||         ::::::   |||||||||:::
Db        366 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 425

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      |||||:::   |||                  :::
Db        426 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 473

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        474 GAGCGCACCATTCCCATCACCCGGGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 527

RESULT 34
AX937703
LOCUS       AX937703                537 bp    DNA     linear   PAT 06-JAN-2004
DEFINITION  Sequence 13 from Patent WO03091266.
ACCESSION   AX937703
VERSION     AX937703.1  GI:40713714
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Quinlan,R.U.
  TITLE     Protein stabilisation
  JOURNAL   Patent: WO 03091266-A 13 06-NOV-2003;
            The University Court of the University of Dundee (GB)
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.3e-45        Length:         537
Score:                  489.00         Matches:        97
Percent Similarity:     73.60%         Conservative:   34
Best Local Similarity:  54.49%         Mismatches:     35
Query Match:            53.38%         Indels:         12
DB:                     6              Gaps:           6

US-10-657-740-1 (1-173) x AX937703 (1-537)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          9 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 68

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||   |||   |||   |||||||   |||
Db         69 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 125

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||      :::
Db        126 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 185

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        186 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 245

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        246 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 305

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::|||||||||||||||||||
Db        306 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 365

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::|||||         ::::::   |||||||||:::
Db        366 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 425

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      |||||:::   |||                  :::
Db        426 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 473

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        474 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 527

RESULT 35
CQ812349
LOCUS       CQ812349                691 bp    DNA     linear   PAT 24-MAY-2004
DEFINITION  Sequence 101 from Patent WO2004038020.
ACCESSION   CQ812349
VERSION     CQ812349.1  GI:47601969
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wittig,R., Poustka,A., Mollenhauer,J. and Schadendorf,D.
  TITLE     Target genes for the diagnosis and treatment of cancer
  JOURNAL   Patent: WO 2004038020-A 101 06-MAY-2004;
            Deutsches Krebsforschungszentrum Stiftung des oeffentliche n Rechts
            (DE)
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    1. .691
                     /note="S45630"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.79e-45        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x CQ812349 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 36
AX330255
LOCUS       AX330255                691 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 764 from Patent WO0194629.
ACCESSION   AX330255
VERSION     AX330255.1  GI:18103233
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 764 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
```

```
Pred. No.:              1.79e-45        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x AX330255 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 37
AX333032
LOCUS       AX333032                691 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 3541 from Patent WO0194629.
ACCESSION   AX333032
VERSION     AX333032.1  GI:18123666
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 3541 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.79e-45        Length:         691
Score:                  489.00          Matches:        97
```

```
Percent Similarity:     73.60%          Conservative:  34
Best Local Similarity:  54.49%          Mismatches:    35
Query Match:            53.38%          Indels:        12
DB:                     6               Gaps:          6

US-10-657-740-1 (1-173) x AX333032 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 38
BD079402
LOCUS       BD079402                691 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Cancer-associated nucleic acids and polypeptides.
ACCESSION   BD079402
VERSION     BD079402.1  GI:22625005
KEYWORDS    JP 2001516009-A/68.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I.,
            OGhare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.
  TITLE     Cancer-associated nucleic acids and polypeptides
  JOURNAL   Patent: JP 2001516009-A 68 25-SEP-2001;
            LUDWIG INSTITUTE FOR CANCER RESEARCH
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001516009-A/68
            PD   25-SEP-2001
            PF   15-JUL-1998 JP 2000503425
            PR   17-JUL-1997 US    08/896164,10-OCT-1997 US    60/061599 PR
            10-OCT-1997 US    60/061765,10-OCT-1997 US    08/948705 PR
            11-OCT-1997 GB    9721697.2,22-JUN-1998 US    09/102322 PI   LLOYD
            J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO  PI   TSENG
            CHEN,
            PI   IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI
              OZLEM TURECI,
            PI   UGUR SAHIN
            PC
            G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC
             A61P35/00,
            PC   C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
            PC   C12N15/00
            CC   Cancer-associated nucleic acids and polypeptides. FH   Key
                   Location/Qualifiers
            FT   source          1. .691
            FT                   /organism='Homo sapiens (human)'.
FEATURES              Location/Qualifiers
     source           1. .691
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.79e-45        Length:        691
Score:                  489.00          Matches:       97
Percent Similarity:     73.60%          Conservative:  34
Best Local Similarity:  54.49%          Mismatches:    35
Query Match:            53.38%          Indels:        12
DB:                     6               Gaps:          6

US-10-657-740-1 (1-173) x BD079402 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 39
S45630
LOCUS       S45630                  691 bp    mRNA    linear   PRI 08-MAY-1993
DEFINITION  alpha B-crystallin=Rosenthal fiber component [human, glioma cell
            line, mRNA, 691 nt].
ACCESSION   S45630
VERSION     S45630.1  GI:256398
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 691)
  AUTHORS   Iwaki,A., Iwaki,T., Goldman,J.E., Ogomori,K., Tateishi,J. and
            Sakaki,Y.
  TITLE     Accumulation of alpha B-crystallin in brains of patients with
            Alexander's disease is not due to an abnormality of the 5'-flanking
            and coding sequence of the genomic DNA
  JOURNAL   Neurosci. Lett. 140 (1), 89-92 (1992)
  MEDLINE   93025869
   PUBMED   1407707
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 114959] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .691
                     /gene="alpha B-crystallin"
     CDS             26. .553
                     /gene="alpha B-crystallin"
                     /note="Rosenthal fiber component"
                     /codon_start=1
                     /product="alpha B-crystallin"
                     /protein_id="AAB23453.1"
                     /db_xref="GI:256399"
                     /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSL
                     SPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                     KHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVSGPERTIP
                     ITREEKPAVTAAPKK"
ORIGIN

Alignment Scores:
Pred. No.:              1.79e-45        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     9               Gaps:           6

US-10-657-740-1 (1-173) x S45630 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg---Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 40
BC007008
LOCUS       BC007008                744 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens crystallin, alpha B, mRNA (cDNA clone MGC:12326
            IMAGE:3933748), complete cds.
ACCESSION   BC007008
VERSION     BC007008.1  GI:13937812
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 744)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 16 Row: j Column: 15
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4503056.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:12326 IMAGE:3933748"
                     /tissue_type="Heart"
                     /clone_lib="NIH_MGC_74"
                     /lab_host="DH10B"
```

```
                /note="Vector: pDNR-LIB"
     gene       1. .744
                /gene="CRYAB"
                /note="synonyms: CRYA2, CTPP2, HSPB5"
                /db_xref="LocusID:1410"
                /db_xref="MIM:123590"
     CDS        47. .574
                /gene="CRYAB"
                /codon_start=1
                /product="crystallin, alpha B"
                /protein_id="AAH07008.1"
                /db_xref="GI:13937813"
                /db_xref="LocusID:1410"
                /db_xref="MIM:123590"
                /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSL
                SPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                KHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVSGPERTIP
                ITREEKPAVTAAPKK"
ORIGIN

Alignment Scores:
Pred. No.:              1.96e-45        Length:         744
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     9               Gaps:           6

US-10-657-740-1 (1-173) x BC007008 (1-744)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         47 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 106

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        107 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 163

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        164 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 223

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        224 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAAAAGGACAGGTTCTCTGTCAACCTG 283

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        284 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 343

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        344 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 403

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::|||||||         ::::::   |||||||||:::
Db        404 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 463

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        464 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 511

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        512 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 565

RESULT 41
AX888028
LOCUS       AX888028                856 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 3891 from Patent EP1033401.
ACCESSION   AX888028
VERSION     AX888028.1  GI:40046779
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 3891 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .856
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             4. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE93953.1"
                     /db_xref="GI:40046780"
                     /translation="MTSAFKLDFLPDMMVXGRLLVPDRINGTANKMNGALDXXDQPDP
                     DAIKMFVGZIPRSWSXKELTSQLTPHTHLATMDIAIHHPWIRRPFFPFHSPSRLFDQF
                     FGEHLLESDLFPTSTSLSPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLDVKHFS
                     PEELKVKVLGDVIEVHGKHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVL
                     TVNGPRKQVSGPERTIPITREEKPAVTAAPKK"
ORIGIN

Alignment Scores:
Pred. No.:              2.35e-45        Length:         856
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x AX888028 (1-856)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        229 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 288

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        289 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 345

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        346 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 405

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        406 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 465

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        466 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 525

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        526 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 585

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::|||||||         ::::::   |||||||||:::
Db        586 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 645

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        646 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 693

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        694 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 747

RESULT 42
```

```
BD027638
LOCUS       BD027638                856 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD027638
VERSION     BD027638.1  GI:22569380
KEYWORDS    JP 2001269182-A/3884.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 856)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 3884 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/3884
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers
            FT   CDS             4. .753.
FEATURES             Location/Qualifiers
     source          1. .856
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.35e-45        Length:         856
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x BD027638 (1-856)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        229 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 288

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        289 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 345

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||   '   :::
Db        346 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 405

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        406 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 465

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        466 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 525

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        526 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 585

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        586 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 645

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                  :::
Db        646 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 693

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        694 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 747

RESULT 43
AX899079
LOCUS       AX899079                893 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 14942 from Patent EP1033401.
ACCESSION   AX899079
VERSION     AX899079.1  GI:40053992
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 14942 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .893
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.48e-45        Length:         893
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x AX899079 (1-893)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        266 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 325

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        326 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 382

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||       :::
Db        383 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 442

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        443 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 502

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        503 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 562

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        563 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 622

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        623 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 682

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                  :::
Db        683 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 730
```

```
Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||||      |||:::||||||
Db        731 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 784

RESULT 44
BD034612
LOCUS       BD034612                 893 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD034612
VERSION     BD034612.1  GI:22576354
KEYWORDS    JP 2001269182-A/10858.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 893)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 10858 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/10858
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US   60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .893
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.48e-45        Length:         893
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x BD034612 (1-893)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          |||||||:::   |||
Db        266 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 325

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||  |||   |||   ||||||  |||
Db        326 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 382

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||       :::
Db        383 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 442

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        443 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 502

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        503 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 562

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
```

```
Db        563 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 622

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||          ::::::   |||||||||:::
Db        623 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 682

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       |||||||:::   |||                 :::
Db        683 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 730

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
Db        731 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 784

RESULT 45
AX899075
LOCUS       AX899075                 911 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 14938 from Patent EP1033401.
ACCESSION   AX899075
VERSION     AX899075.1  GI:40053988
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 14938 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.54e-45        Length:         911
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x AX899075 (1-911)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          |||||||:::   |||
Db        284 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 343

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||  |||   |||   ||||||  |||
Db        344 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 400

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||       :::
Db        401 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 460

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        461 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 520

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        521 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 580

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        581 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 640

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||          ::::::   |||||||||:::
```

```
Db        641 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 700

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||              :::
Db        701 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 748

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              |||||| ||||||::::::||||||||||||||||      |||:::|||||| 
Db        749 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 802

RESULT 46
BD034608
LOCUS       BD034608                911 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD034608
VERSION     BD034608.1  GI:22576350
KEYWORDS    JP 2001269182-A/10854.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 10854 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/10854
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .911
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.54e-45        Length:         911
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x BD034608 (1-911)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        284 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 343

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        344 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 400

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        401 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 460

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        461 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 520

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        521 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 580

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||   ||||||||||:::|||||:::||||||||||||||||||||||
Db        581 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 640

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||      ::::::   |||||||||:::
Db        641 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 700

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||              :::
Db        701 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 748

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              |||||| ||||||::::::||||||||||||||||      |||:::|||||| 
Db        749 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 802

RESULT 47
AX899073
LOCUS       AX899073                927 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 14936 from Patent EP1033401.
ACCESSION   AX899073
VERSION     AX899073.1  GI:40053986
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 14936 06-SEP-2000;
            Genset (FR)
FEATURES              Location/Qualifiers
     source           1. .927
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.6e-45         Length:         927
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x AX899073 (1-927)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        300 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 359

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        360 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 416

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        417 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 476

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        477 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 536

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        537 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 596

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
```

```
                ::::|||||||||||||   ||||||||||||:::|||||||:::|||||||||||||||||||
Db        597 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 656

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        657 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 716

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        717 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 764

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        765 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 818

RESULT 48
BD034606
LOCUS       BD034606                927 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD034606
VERSION     BD034606.1  GI:22576348
KEYWORDS    JP 2001269182-A/10852.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 927)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 10852 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/10852
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .927
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.6e-45       Length:        927
Score:                  489.00        Matches:       97
Percent Similarity:     73.60%        Conservative:  34
Best Local Similarity:  54.49%        Mismatches:    35
Query Match:            53.38%        Indels:        12
DB:                     6             Gaps:          6

US-10-657-740-1 (1-173) x BD034606 (1-927)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        300 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 359

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        360 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 416

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||       |||       :::
Db        417 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 476

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              |||:::|||:::||||||:::|||   ::::::|||:::|||   :::    |||
Db        477 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 536

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        537 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 596

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        597 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 656

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        657 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 716

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                 :::
Db        717 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 764

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        765 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 818

RESULT 49
AX888027
LOCUS       AX888027                942 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 3890 from Patent EP1033401.
ACCESSION   AX888027
VERSION     AX888027.1  GI:40046777
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 3890 06-SEP-2000;
            Genset (FR)
FEATURES              Location/Qualifiers
     source           1. .942
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
     CDS              315. .842
                      /note="unnamed protein product"
                      /codon_start=1
                      /protein_id="CAE93952.1"
                      /db_xref="GI:40046778"
                      /translation="MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSL
                      SPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHG
                      KHEERQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVSGPERTIP
                      ITREEKPAVTAAPKK"
ORIGIN

Alignment Scores:
Pred. No.:              2.65e-45      Length:        942
Score:                  489.00        Matches:       97
Percent Similarity:     73.60%        Conservative:  34
Best Local Similarity:  54.49%        Mismatches:    35
Query Match:            53.38%        Indels:        12
DB:                     6             Gaps:          6

US-10-657-740-1 (1-173) x AX888027 (1-942)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        315 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 374

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        375 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 431
```

```
Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
Db        432 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 491

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
Db        492 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 551

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
Db        552 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 611

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
Db        612 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 671

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
Db        672 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 731

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
Db        732 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 779

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
Db        780 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 833
```

RESULT 50
BD027637
LOCUS BD027637 942 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD027637
VERSION BD027637.1 GI:22569379
KEYWORDS JP 2001269182-A/3883.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 942)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 3883 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/3883
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS 315. .839.
FEATURES Location/Qualifiers
source 1. .942
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	2.65e-45	Length:	942
Score:	489.00	Matches:	97
Percent Similarity:	73.60%	Conservative:	34
Best Local Similarity:	54.49%	Mismatches:	35
Query Match:	53.38%	Indels:	12
DB:	6	Gaps:	6

US-10-657-740-1 (1-173) x BD027637 (1-942)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
Db        315 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 374

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
Db        375 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 431

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
Db        432 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 491

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
Db        492 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 551

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
Db        552 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 611

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
Db        612 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 671

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
Db        672 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 731

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
Db        732 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 779

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
Db        780 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 833
```

Search completed: May 30, 2005, 07:14:52
Job time : 4226.45 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 29, 2005, 22:54:38 ; Search time 526.598 Seconds
(without alignments)
1944.776 Million cell updates/sec

Title: US-10-657-740-1
Perfect score: 916
Sequence: 1 MDVTIQHPWFKRTLGPFYPS.....HAERAIPVSREEKPTSAPSS 173

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10657740/runat_27052005_165252_3305/app_query.fasta_1.590
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=150 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10657740_@CGN_1_1_796_@runat_27052005_165252_3305 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	871	95.1	1056	10	ADB52521	Adb52521 Primary r
2	871	95.1	1056	10	ADF30549	Adf30549 Rat angio
3	849.5	92.7	1271	10	ADF30547	Adf30547 Rat angio
4	702	76.6	531	8	ABX12062	Abx12062 Human alp
5	639	69.8	372	13	ADQ78288	Adq78288 Human N-t
6	493	53.8	548	6	ABZ35138	Abz35138 Human gen
7	489	53.4	528	13	ADR46388	Adr46388 Human len
8	489	53.4	537	10	ADE75374	Ade75374 Human mut
9	489	53.4	537	10	ADE75375	Ade75375 Human wil
10	489	53.4	691	2	AAX39668	Aax39668 Renal can
11	489	53.4	691	6	ABL65204	Abl65204 Lung canc
12	489	53.4	691	6	ABL62427	Abl62427 Colon ade
13	489	53.4	691	6	ABT10966	Abt10966 Human bre
14	489	53.4	691	6	ABK64671	Abk64671 Human ben
15	489	53.4	691	10	ADF09686	Adf09686 Human cry
16	489	53.4	691	12	ADN05800	Adn05800 Antipsori
17	489	53.4	691	13	ADR24825	Adr24825 Breast ca
18	489	53.4	856	3	AAC03893	Aac03893 Human sec
19	489	53.4	893	3	AAC10867	Aac10867 Human sec
20	489	53.4	911	3	AAC10863	Aac10863 Human sec
21	489	53.4	927	3	AAC10861	Aac10861 Human sec
22	489	53.4	942	3	AAC03892	Aac03892 Human sec
c 23	489	53.4	1036	2	AAZ33574	Aaz33574 Human bre
24	485	52.9	528	10	ADB52642	Adb52642 Primary r
25	485	52.9	528	10	ABT41872	Abt41872 Toxicity
26	485	52.9	1247	6	ABK63512	Abk63512 Rat seque
27	485	52.9	1247	12	ADP72683	Adp72683 Renal tox
28	457.5	49.9	913	3	AAC10862	Aac10862 Human sec
29	418	45.6	449	8	ABX39112	Abx39112 Bovine ES
30	394	43.0	396	8	ABX47457	Abx47457 Bovine ES
31	367	40.1	211	12	ACH93386	Ach93386 Human gen
32	367	40.1	573	12	ACH79686	Ach79686 Human gen
33	364.5	39.8	411	11	ACN89210	Acn89210 Breast ca
34	361	39.4	450	9	ACH17737	Ach17737 Human adu
35	347	37.9	579	12	ACH67866	Ach67866 Human gen
36	340	37.1	480	10	ADH76210	Adh76210 Human hea
37	340	37.1	575	6	ABQ61117	Abq61117 Skeletal
38	340	37.1	1433	5	ADM19230	Adm19230 Novel hum
39	340	37.1	1488	2	AAZ42230	Aaz42230 Human nor
c 40	340	37.1	1506	5	ADM19501	Adm19501 Novel hum
41	340	37.1	1771	3	AAC76681	Aac76681 Human ORF
42	340	37.1	1820	12	ADQ22255	Adq22255 Human sof
43	338	36.9	212	12	ACH81569	Ach81569 Human gen
44	338	36.9	421	9	ACH17560	Ach17560 Human adu
45	336	36.7	486	10	ADH76202	Adh76202 Rat heat
46	336	36.7	1310	10	ABT41710	Abt41710 Toxicity
47	332.5	36.3	380	8	ABX38978	Abx38978 Bovine ES
48	330	36.0	695	3	AAC03894	Aac03894 Human sec
49	311.5	34.0	405	9	ACH16863	Ach16863 Human adu
50	308.5	33.7	719	4	ABL07843	Abl07843 Drosophil
51	296.5	32.4	387	12	ADQ17563	Adq17563 Human sof
52	295.5	32.3	543	12	ADL11837	Adl11837 Cat flea
53	291	31.8	349	6	ABK87386	Abk87386 Target mo
54	289.5	31.6	615	12	ADN31647	Adn31647 Human Hsp
55	289.5	31.6	764	12	ADM94741	Adm94741 Human hea
56	289.5	31.6	789	11	ADI31974	Adi31974 Human cDN
57	289.5	31.6	847	10	ADG10693	Adg10693 Human STA
58	289.5	31.6	847	13	ADQ86169	Adq86169 Human tum
59	289.5	31.6	865	6	ABQ60780	Abq60780 Human HSB
60	289.5	31.6	865	13	ADR24765	Adr24765 Breast ca
61	289.5	31.6	865	13	ACN38656	Acn38656 Tumour-as
62	289.5	31.6	865	13	ADP23292	Adp23292 PRO polyp
63	289.5	31.6	1231	6	ABN97370	Abn97370 Gene #386
64	289.5	31.6	1231	10	ADD70998	Add70998 Human hea
65	289.5	31.6	1380	3	AAA93441	Aaa93441 GFP-HSP27
66	289.5	31.6	1380	6	ABS71560	Abs71560 DNA encod
67	289	31.6	787	6	ABK63521	Abk63521 Rat seque
68	289	31.6	787	12	ADP72725	Adp72725 Renal tox
69	289	31.6	787	13	ADT66606	Adt66606 Rat heat
c 70	288.5	31.5	563	11	ACN87226	Acn87226 Breast ca
71	288.5	31.5	787	10	ADB58686	Adb58686 Toxicity-
72	288.5	31.5	787	10	ABT42261	Abt42261 Toxicity
73	288	31.4	599	6	ABV94670	Abv94670 Human pan
74	288	31.4	618	2	AAQ48718	Aaq48718 Encodes p
75	284	31.0	560	10	ADK12087	Adk12087 Breast ca
76	283	30.9	1379	3	AAC66073	Aac66073 E. coli e
77	283	30.9	1379	4	AAF77805	Aaf77805 Fragment
78	283	30.9	1379	4	AAF61192	Aaf61192 Expressio

	79	283	30.9	1379	5	AAD02211 Aad02211 Partial p
	80	276	30.1	503	6	ABL99507 Abl99507 Target ca
	81	275	30.0	558	10	ADK12085 Adk12085 Breast ca
c	82	266.5	29.1	604	6	ABK63074 Abk63074 Rat seque
c	83	266.5	29.1	604	12	ADP72259 Adp72259 Renal tox
c	84	266.5	29.1	2874	4	ABL07842 Abl07842 Drosophil
c	85	266.5	29.1	3119	4	ABL09418 Abl09418 Drosophil
	86	266	29.0	410	6	ABS70907 Abs70907 Deer cDNA
	87	265.5	29.0	909	2	AAQ03969 Aaq03969 Clone P25
	88	263.5	28.8	722	13	ADT66608 Adt66608 Murine he
	89	259.5	28.3	654	5	AAS42494 Aas42494 Human cDN
	90	257	28.1	1700	4	AAS56378 Aas56378 Human cDN
	91	256.5	28.0	755	2	AAQ03970 Aaq03970 Clone P25
c	92	251	27.4	376	4	AAL19387 Aal19387 Human bre
	93	251	27.4	549	3	AAC93846 Aac93846 Cat flea
	94	247	27.0	550	12	ADL11752 Adl11752 Cat flea
	95	241.5	26.4	708	5	ADL40880 Adl40880 Human ova
	96	234.5	25.6	614	5	ADI75644 Adi75644 Human ova
	97	234.5	25.6	614	5	ADI69298 Adi69298 Human ova
	98	233.5	25.5	954	5	ADL62061 Adl62061 Human ova
	99	225.5	24.6	542	12	ADL11704 Adl11704 Cat flea
	100	223.5	24.4	500	9	ACH34922 Ach34922 Human end
	101	223	24.3	467	6	ABV94661 Abv94661 Human pan
	102	220.5	24.1	482	9	ACH17221 Ach17221 Human adu
	103	215	23.5	870	4	ABL07721 Abl07721 Drosophil
c	104	215	23.5	2870	4	ABL07720 Abl07720 Drosophil
c	105	212	23.1	49634	6	ABL68647 Abl68647 Kidney ca
	106	205.5	22.4	212	8	ABT33639 Abt33639 Anticance
	107	202.5	22.1	937	4	ABL07315 Abl07315 Drosophil
	108	202.5	22.1	2651	4	ABL07714 Abl07714 Drosophil
	109	202.5	22.1	2937	4	ABL07314 Abl07314 Drosophil
c	110	198.5	21.7	396	6	ABL63575 Abl63575 Breast ca
c	111	198.5	21.7	396	6	ABL63995 Abl63995 Breast ca
	112	196	21.4	1206	4	ABL07725 Abl07725 Drosophil
c	113	196	21.4	3206	4	ABL07724 Abl07724 Drosophil
	114	193	21.1	600	4	ABL07321 Abl07321 Drosophil
	115	193	21.1	2600	4	ABL07320 Abl07320 Drosophil
	116	192	21.0	6806	9	ACF25379 Acf25379 Rat alpha
	117	191.5	20.9	445	9	ACH23297 Ach23297 Human adu
	118	188	20.5	4206	6	ABZ35176 Abz35176 Human gen
	119	188	20.5	4206	10	ADE84874 Ade84874 Farnesyl
	120	187	20.4	2632	8	ABX63312 Abx63312 Human cDN
	121	182.5	19.9	393	3	AAC06428 Aac06428 Human sec
	122	182.5	19.9	487	13	ADS54436 Ads54436 Bacterial
	123	182.5	19.9	487	13	ADS54437 Ads54437 Bacterial
c	124	180.5	19.7	565	11	ACN79216 Acn79216 Breast ca
c	125	179	19.5	218	4	AAL09000 Aal09000 Human bre
	126	179	19.5	1811	12	ADJ75836 Adj75836 Marker ge
	127	174.5	19.1	1367	4	ABL07295 Abl07295 Drosophil
	128	174.5	19.1	3367	4	ABL07294 Abl07294 Drosophil
c	129	174	19.0	510	6	ABQ14062 Abq14062 Oligonucl
	130	174	19.0	510	6	ABQ14063 Abq14063 Oligonucl
	131	172.5	18.8	1495	10	ADI22459 Adi22459 Rat liver
	132	171.5	18.7	1627	2	AAZ21959 Aaz21959 cDNA enco
	133	171.5	18.7	1627	9	ACH04117 Ach04117 Human cDN
	134	171.5	18.7	1843	4	AAF74328 Aaf74328 Human H11
	135	171.5	18.7	1843	4	AAF74307 Aaf74307 Human H11
	136	171.5	18.7	1854	5	ABX71326 Abx71326 Human met
	137	171.5	18.7	2004	5	ADM19232 Adm19232 Novel hum
	138	171.5	18.7	2010	12	ADJ74941 Adj74941 Marker ge
	139	171.5	18.7	2010	13	ADR25621 Adr25621 Breast ca
c	140	171.5	18.7	2012	4	AAK52551 Aak52551 Human pol
	141	171.5	18.7	2048	4	AAK51567 Aak51567 Human pol
	142	165	18.0	651	4	ABL07715 Abl07715 Drosophil
c	143	165	18.0	2651	4	ABL07714 Abl07714 Drosophil
c	144	165	18.0	2937	4	ABL07314 Abl07314 Drosophil
	145	163.5	17.8	2036	10	ADI22637 Adi22637 Human liv
	146	163	17.8	509	3	AAC01050 Aac01050 Human sec
	147	163	17.8	546	3	AAC01053 Aac01053 Human sec
	148	163	17.8	564	3	AAC01049 Aac01049 Human sec
	149	163	17.8	580	3	AAC01051 Aac01051 Human sec
	150	163	17.8	595	3	AAC01048 Aac01048 Human sec

ALIGNMENTS

RESULT 1
ADB52521
ID ADB52521 standard; DNA; 1056 BP.
XX
AC ADB52521;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3063.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3063; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX

```
SQ   Sequence 1056 BP; 199 A; 333 C; 269 G; 255 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.07e-102      Length:         1056
Score:                  871.00         Matches:        164
Percent Similarity:     97.69%         Conservative:   5
Best Local Similarity:  94.80%         Mismatches:     4
Query Match:            95.09%         Indels:         0
DB:                     10             Gaps:           0

US-10-657-740-1 (1-173) x ADB52521 (1-1056)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         13 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 72

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 132

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 192

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTGAAGCACTTC 252

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        253 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCATGGCAAACAC 312

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 372

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGCATGCTGACC 432

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        433 TTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGGGCCATTCCC 492

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::|||||||||||||||
Db        493 GTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 531

RESULT 2
ADF30549
ID   ADF30549 standard; cDNA; 1056 BP.
XX
AC   ADF30549;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Rat angiogenesis modulating protein cDNA #29.
XX
KW   ss; gene; rat; angiogenesis; angiogenesis modulating protein;
KW   retinal neovascularisation; choroidal neovascularisation;
KW   chronic inflammation; myocardial ischaemia; stroke;
KW   coronary artery disease; peripheral vascular disease.
XX
OS   Rattus norvegicus.
XX
PN   US2003162706-A1.
XX
PD   28-AUG-2003.
XX
PF   10-DEC-2002; 2002US-00316253.
XX
PR   08-FEB-2002; 2002US-0355295P.
PR   26-JUN-2002; 2002US-0391758P.
XX
PA   (PROC ) PROCTER & GAMBLE CO.
XX
PI   Peters KG,  Thompson LJ,  Wang F,  Greis KD;
XX
DR   WPI; 2003-711557/67.
DR   P-PSDB; ADF30550.
XX
PT   Treating angiogenesis-mediated disorder, e.g., retinal or choroidal
PT   neovascularization or diseases associated with chronic inflammation,
PT   myocardial ischemia, stroke, coronary artery disease or peripheral
PT   vascular disease.
XX
PS   Disclosure; SEQ ID NO 112; 26pp; English.
XX
CC   The invention relates to a method of treating an angiogenesis-mediated
CC   disorder in a subject. The method is useful for treating angiogenesis-
CC   mediated disorder, e.g., retinal or choroidal neovascularisation or
CC   diseases associated with chronic inflammation, myocardial ischaemia,
CC   stroke, coronary artery disease or peripheral vascular disease. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 1056 BP; 199 A; 333 C; 269 G; 255 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.07e-102      Length:         1056
Score:                  871.00         Matches:        164
Percent Similarity:     97.69%         Conservative:   5
Best Local Similarity:  94.80%         Mismatches:     4
Query Match:            95.09%         Indels:         0
DB:                     10             Gaps:           0

US-10-657-740-1 (1-173) x ADF30549 (1-1056)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         13 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 72

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 132

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 192

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTGAAGCACTTC 252

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        253 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCATGGCAAACAC 312

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 372

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGCATGCTGACC 432

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        433 TTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGGGCCATTCCC 492

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||:::|||||||||||||||
Db        493 GTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 531

RESULT 3
ADF30547
ID   ADF30547 standard; cDNA; 1271 BP.
```

```
XX
AC   ADF30547;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Rat angiogenesis modulating protein cDNA #28.
XX
KW   ss; gene; rat; angiogenesis; angiogenesis modulating protein;
KW   retinal neovascularisation; choroidal neovascularisation;
KW   chronic inflammation; myocardial ischaemia; stroke;
KW   coronary artery disease; peripheral vascular disease.
XX
OS   Rattus norvegicus.
XX
PN   US2003162706-A1.
XX
PD   28-AUG-2003.
XX
PF   10-DEC-2002; 2002US-00316253.
XX
PR   08-FEB-2002; 2002US-0355295P.
PR   26-JUN-2002; 2002US-0391758P.
XX
PA   (PROC ) PROCTER & GAMBLE CO.
XX
PI   Peters KG,  Thompson LJ,  Wang F,  Greis KD;
XX
DR   WPI; 2003-711557/67.
DR   P-PSDB; ADF30548.
XX
PT   Treating angiogenesis-mediated disorder, e.g., retinal or choroidal
PT   neovascularization or diseases associated with chronic inflammation,
PT   myocardial ischemia, stroke, coronary artery disease or peripheral
PT   vascular disease.
XX
PS   Disclosure; SEQ ID NO 110; 26pp; English.
XX
CC   The invention relates to a method of treating an angiogenesis-mediated
CC   disorder in a subject. The method is useful for treating angiogenesis-
CC   mediated disorder, e.g., retinal or choroidal neovascularisation or
CC   diseases associated with chronic inflammation, myocardial ischaemia,
CC   stroke, coronary artery disease or peripheral vascular disease. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 1271 BP; 263 A; 384 C; 315 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.76e-99        Length:         1271
Score:                    849.50          Matches:        164
Percent Similarity:       86.22%          Conservative:   5
Best Local Similarity:    83.67%          Mismatches:     4
Query Match:              92.74%          Indels:         23
DB:                       10              Gaps:           1

US-10-657-740-1 (1-173) x ADF30547 (1-1271)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        159 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 218

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 278

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 338

Qy         61 IleSerGlu------------------------------------------------- 63
              |||||||||
Db        339 ATCTCTGAGCTCATGACCCATATGTGGTTTGTAATGCACCAACCACATGCTGGAAACCCC 398

Qy         64 ------------------ValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                                ||||||||||||||||||||||||||||||||||||||||||
Db        399 AAGAACAACCCCGGCAAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTG 458

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||||||||||||||||||||||||    :::||||||||||||||||||
Db        459 AAGCACTTCTCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCAT 518

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 GGCAAACACAACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGC 578

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 TACCGTCTGCCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGC 638

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              ||||||||||||   |||||||||:::|||:::||||||||||||   |||:::||||||
Db        639 ATGCTGACCTTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGG 698

Qy        158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||||||||||||:::|||||||||||||||
Db        699 GCCATTCCCGTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 746

RESULT 4
ABX12062
ID   ABX12062 standard; DNA; 531 BP.
XX
AC   ABX12062;
XX
DT   16-MAY-2003  (first entry)
XX
DE   Human alpha BNAC crystallin chimera DNA.
XX
KW   Human; ds; gene; chimera; alpha BNAC crystallin; protein shelf life;
KW   protein aggregation; accessible hydrophobic region increase; mutant;
KW   larger size oligomer formation; intersubunit interaction increase;
KW   larger aggregate formation; larger porous oligomer formation;
KW   increased ellipticity; less solvent accessible tryptophan;
KW   increased chaperone-like activity; alpha A crystallin;
KW   alpha B crystallin.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .531
FT                   /*tag=  a
FT                   /product= "Alpha BNAC crystallin"
FT                   /partial
FT                   /note= "No stop codon given"
FT   misc_feature    1. .247
FT                   /*tag=  b
FT                   /note= "Derived from DNA sequence encoding N-terminal
FT                   sequence of alpha B crystallin"
FT   misc_feature    248. .531
FT                   /*tag=  c
FT                   /note= "Derived from DNA sequence encoding N-terminal
FT                   sequence of alpha A crystallin"
XX
PN   US2002177192-A1.
XX
PD   28-NOV-2002.
XX
PF   26-MAR-2002; 2002US-00105427.
XX
PR   28-MAR-2001; 2001US-0279223P.
XX
PA   (KUMA/) KUMAR L V S.
PA   (RAOC/) RAO C M.
XX
PI   Kumar LVS,  Rao CM;
XX
DR   WPI; 2003-298776/29.
```

```
DR   P-PSDB; ABG76084.
XX
PT   New chimera alpha BNAC nucleic acid, useful for preventing aggregation of
PT   proteins and also for increasing shelf life of proteins of pharmaceutical
PT   value.
XX
PS   Claim 1; Fig 9; 17pp; English.
XX
CC   The invention relates to a chimera alpha BNAC polynucleotide that encodes
CC   a chimeric alpha BNAC polypeptide. The polypeptide is useful for
CC   preventing protein aggregation. The polypeptide is also useful for
CC   increasing the shelf life of proteins of pharmaceutical value. The
CC   polypeptide shows an increase in accessible hydrophobic regions, forms
CC   larger size oligomers, shows an increase in intersubunit interaction,
CC   forms larger aggregates, forms larger porous oligomers and shows
CC   increased ellipticity as compared to eye lens crystallins alpha A and
CC   alpha B. The tryptophan residues in the polypeptide are less solvent
CC   accessible as compared to those of eye lens crystallins alpha A and alpha
CC   B. The polypeptide shows extraordinarily high chaperone-like activity
CC   ranging between 3 - 6 times that of the eye lens crystallins alpha A and
CC   alpha B. The present sequence represents the chimera DNA that encodes
CC   human alpha BNAC crystallin
XX
SQ   Sequence 531 BP; 90 A; 198 C; 130 G; 113 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.38e-80      Length:          531
Score:                  702.00        Matches:         137
Percent Similarity:     85.96%        Conservative:    16
Best Local Similarity:  76.97%        Mismatches:      19
Query Match:            76.64%        Indels:          6
DB:                     8             Gaps:            4

US-10-657-740-1 (1-173) x ABX12062 (1-531)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          |||||||:::   |||
Db          1 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||::::::::::|||||||:::|||   :::      |||      |||      :::
Db        118 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        178 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 GATGTGAAGCACTTCTCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ATCCACGGAAAGCACAACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCAC 357

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 CGCCGCTACCGCCTGCCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCC 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GATGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCC 477

Qy        156 GluArgAlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 GAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 531

RESULT 5
ADQ78288
ID   ADQ78288 standard; cDNA; 372 BP.
XX
AC   ADQ78288;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Human N-terminal truncated alpha-crystallin DNA.
XX
KW   alpha-crystallin; enhanced protein expression;
KW   enhanced protein secretion; protein aggregation; heat tolerance;
KW   elevated temperature; human; ds; gene.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .372
FT                   /*tag=  a
FT                   /partial
FT                   /product= "N-terminal truncated alpha-crystallin"
FT                   /note= "No start codon given"
XX
PN   US2004157289-A1.
XX
PD   12-AUG-2004.
XX
PF   08-SEP-2003; 2003US-00657740.
XX
PR   06-SEP-2002; 2002US-0408680P.
XX
PA   (SALE/) SALERNO J C.
PA   (HANN/) HANNA M.
PA   (KORE/) KORETZ J F.
PA   (CRON/) CRONE D.
PA   (SMIT/) SMITH S M E.
XX
PI   Salerno JC,  Hanna M,  Koretz JF,  Crone D,  Smith SME;
XX
DR   WPI; 2004-580268/56.
DR   P-PSDB; ADQ78289.
XX
PT   New truncated alpha-crystallin polypeptide derived from a wild-type alpha
PT   -crystallin protein, useful for enhancing protein (e.g. insulin or
PT   alcohol dehydrogenase) expression or secretion and for preventing protein
PT   aggregation.
XX
PS   Claim 16; SEQ ID NO 2; 33pp; English.
XX
CC   The invention relates to a truncated alpha-crystallin polypeptide derived
CC   from a wild-type alpha-crystallin protein, where the truncated
CC   polypeptide lacks an N-terminal sequence present in the wild-type
CC   protein. The composition and methods are useful for enhancing protein
CC   (e.g. insulin or alcohol dehydrogenase) expression or secretion and for
CC   preventing protein aggregation. These may also be used for creating a
CC   thermophilic host that tolerates elevated temperatures. The present
CC   sequence represents human N-terminal truncated alpha-crystallin DNA.
XX
SQ   Sequence 372 BP; 66 A; 136 C; 98 G; 72 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.09e-72      Length:          372
Score:                  639.00        Matches:         123
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            69.76%        Indels:          0
DB:                     13            Gaps:            0

US-10-657-740-1 (1-173) x ADQ78288 (1-372)

Qy         51 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 70
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGGACAAG 60

Qy         71 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 90
```

```
Db         61 TTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGACCTCACCGTGAAGGTGCAG 120

Qy         91 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 110

Db        121 GACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAGGACGACCACGGCTACATT 180

Qy        111 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 130

Db        181 TCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTGGACCAGTCGGCCCTCTCT 240

Qy        131 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 150

Db        241 TGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTG 300

Qy        151 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 170

Db        301 GATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAGCCCACCTCGGCT 360

Qy        171 ProSerSer 173

Db        361 CCCTCGTCC 369
```

RESULT 6
ABZ35138
ID ABZ35138 standard; cDNA; 548 BP.
XX
AC ABZ35138;
XX
DT 05-FEB-2003 (first entry)
XX
DE Human gene expression profile polynucleotide SEQ ID NO 250.
XX
KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW gene expression; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200274979-A2.
XX
PD 26-SEP-2002.
XX
PF 20-MAR-2002; 2002WO-US008456.
XX
PR 20-MAR-2001; 2001US-0276947P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI Wan J, Wang Y;
XX
DR WPI; 2002-740862/80.
XX
PT New gene expression profile generated from primary, endothelial,
PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer.
XX
PS Claim 10; Page 429; 850pp; English.
XX
CC The invention relates to a gene expression profile comprising one or more
CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes, DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC identifications and in identifying promising antibiotics, antiviral or
CC antifungal agents
XX
SQ Sequence 548 BP; 120 A; 178 C; 127 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.33e-53	Length:	548
Score:	493.00	Matches:	98
Percent Similarity:	74.16%	Conservative:	34
Best Local Similarity:	55.06%	Mismatches:	34
Query Match:	53.82%	Indels:	12
DB:	6	Gaps:	6

US-10-657-740-1 (1-173) x ABZ35138 (1-548)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19

Db         21 ATGGACATCGCTATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTTCACTCGCCC 80

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39

Db         81 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTCTTCCCA--- 137

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55

Db        138 ACTTCTACTTCCCTGAGCCCCTTCTATCTTCGGCCACCCTCATTCCTGCGGGCACCCAGC 197

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75

Db        198 TGGATTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 257

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95

Db        258 GATGTGAAGCACTTCTCCCCAGAGGAGCTCAAGGTCAAAGTGTTGGGTGATGTGATTGAG 317

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115

Db        318 GTGCACGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 377

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135

Db        378 AGGAAATACCGGATCCCAGCTGATGTGGACCCTCTCACCATTACTTCATCCCTGTCATCT 437

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155

Db        438 GATGGGGTCCTCACTGTGAATGGACCAAGGAAGCAA-----------GCCCCTGGCCCA 485

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171

Db        486 GAGCGCACCATTCCCATAACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCC 539
```

RESULT 7
ADR46388
ID ADR46388 standard; DNA; 528 BP.
XX
AC ADR46388;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human lens crystalline protein coding sequence.
XX
KW cytostatic; immunotoxin; cancer; mitochondrial malate dehydrogenase;
KW enzyme; human; MDH; ds; gene; lens crystalline protein.
XX

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .528
FT                   /*tag=  a
FT                   /product= "lens crystalline protein"
XX
PN   WO2004070012-A2.
XX
PD   19-AUG-2004.
XX
PF   02-FEB-2004; 2004WO-US002974.
XX
PR   02-FEB-2003; 2003US-00444191.
PR   03-FEB-2003; 2003US-0444191P.
PR   08-APR-2003; 2003US-0460855P.
XX
PA   (PALO-) PALO ALTO INST MOLECULAR MEDICINE.
XX
PI   Wright SC,  Larrick JW,  Nock SR,  Wilson DS;
XX
DR   WPI; 2004-604434/58.
DR   P-PSDB; ADR46387.
XX
PT   New compositions comprising proteins and encoding nucleic acids having a
PT   DNA nuclease or cell killing activity and are operably linked to cancer
PT   cell binding antibodies or growth factors, useful for treating cancer.
XX
PS   Disclosure; SEQ ID NO 35; 225pp; English.
XX
CC   The present invention relates to a composition comprising an isolated
CC   amino acid sequence that comprises a portion of human mitochondrial
CC   malate dehydrogenase protein (MDH), particularly the minimun activator of
CC   DNA fragmentation and activator of DNA fragmentation sequences. The
CC   composition is useful for treating cancer chosen from liver cancer,
CC   gastric cancer, head cancer, neck cancer, lung cancer, breast cancer,
CC   prostate cancer, cervical cancer, pancreatic cancer, colon cancer,
CC   ovarian cancer, stomach cancer, oesophagus cancer, mouth cancer, tongue
CC   cancer, gum cancer, skin cancer, muscle cancer, heart cancer, bronchial
CC   cancer, cartilage cancer, bone cancer, testis cancer, kidney cancer,
CC   endometrium cancer, uterus cancer, bladder cancer, bone marrow cancer,
CC   lymphoma cancer, spleen cancer, thymus cancer, thyroid cancer, brain
CC   cancer, neuron cancer, gall bladder cancer, ocular cancer, joint cancer,
CC   glioblastoma, mesothelioma, lymphoma, leukaemia, melanoma, squamous cell
CC   carcinoma, osteosarcoma, and Kaposi's sarcoma. The present sequence is a
CC   coding sequence shown in the exemplification of the invention.
XX
SQ   Sequence 528 BP; 115 A; 167 C; 123 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               4.11e-53      Length:          528
Score:                   489.00        Matches:         97
Percent Similarity:      73.60%        Conservative:    34
Best Local Similarity:   54.49%        Mismatches:      35
Query Match:             53.38%        Indels:          12
DB:                      13            Gaps:            6

US-10-657-740-1 (1-173) x ADR46388 (1-528)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||        ||||||:::   |||
Db          1 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        118 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||

Db        178 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        238 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        298 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 357

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||         ::::::   |||||||||:::
Db        358 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||              :::
Db        418 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 465

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 519

RESULT 8
ADE75374
ID   ADE75374 standard; DNA; 537 BP.
XX
AC   ADE75374;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human mutant alphaB-crystallin fragment-encoding DNA.
XX
KW   Hybrid protein chaperone; protein stabilisation; heat shock protein;
KW   sHSP family; protein aggregation inhibition; cell death inhibition;
KW   genome stability pathway inhibition; protein denaturation identification;
KW   protein conformation related disease; cardiomyopathy; cataract;
KW   neurodegenerative disease; cardiant; ophthalmological; neuroprotective;
KW   gene therapy; alphaB-crystallin; wild-type; human; cassette mutagenesis;
KW   mutant; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    493. .498
FT                   /*tag=  a
FT                   /note= "Unique AvaI site for casette mutagenesis"
FT   mutation        replace(497,T)
FT                   /*tag=  b
FT                   /note= "Mutagenesis to G from T at this site generates a
FT                   unique AvaI site between bases 493-498"
XX
PN   WO2003091266-A2.
XX
PD   06-NOV-2003.
XX
PF   23-APR-2003; 2003WO-GB001721.
XX
PR   23-APR-2002; 2002GB-00009334.
XX
PA   (UYDU-) UNIV DUNDEE.
XX
PI   Quinlan R;
XX
DR   WPI; 2003-865571/80.
XX
PT   New hybrid protein chaperone (e.g. heat shock protein) useful for
PT   stabilizing proteins and/or protein activities, or as an agent to prevent
PT   protein aggregation, or for treating diseases involving altered protein
PT   conformations.
XX
PS   Disclosure; Fig 12; 45pp; English.
XX
```

```
CC   The invention relates to a hybrid protein chaperone for stabilising
CC   proteins and/or protein activities. Protein chaperones (also known as
CC   heat shock proteins) are divided into 4 families on the basis of their
CC   primary sequence and chaperone properties: HSP90, HSP70, HSP60 and sHSP
CC   (small heat shock protein). The invention is based upon the finding that
CC   among the sHSP family, which have a general structure of a cenral domain
CC   (called the alpha-crystallin domain) flanked by N and C-terminal regions,
CC   replacement of one or more regions of an sHSP with the corresponding
CC   region from a second sHSP can improve the activity compared to native
CC   sHSPs. In a particualr embodiment of the invention, the hybrid chaperone
CC   is a hybrid sHSP designated alphaB-HSP27 comprising the N-terminus and
CC   central portion of alphaB-crystallin and the C-terminal tail of HSP27.
CC   However, the hybrid protein chaperones of the invention can comprise
CC   regions from HSP90, HSP70 and HSP60 families as well as from the sHSP
CC   family. The invention also encompasses methods for stabilising proteins
CC   such as enzymes, therapuetic proteins, diagnostic proteins, reporter
CC   proteins or antibodies, their fragments or conjugates in an aqueous
CC   solution using hybrid protein chaperones; stabilised protein formulation
CC   comprising at least one protein associated with the above hybrid protein
CC   chaperone; nucleic acids and vectors encoding a hybrid protein chaperone;
CC   and the coexpression of a recombinant protein of interest and a hybrid
CC   protein chaperone. The hybrid protein chaperones are useful as agents to
CC   prevent protein aggregation, as inhibitors of cell death and genome
CC   stability pathways, for identification of proteins that are in the
CC   process of unfolding, for the treatment of diseases involving altered
CC   protein conformations (e.g., cardiomyopathies, cataracts or
CC   neurodegenerative disease), or for the manufacture of a medicament for
CC   the treatment of such diseases. The present sequence represents a mutant
CC   human alphaB-crystallin DNA sequence in which the G at position 497 was
CC   mutated from T (see ADE75375) to generate a unique AvaI restriction site
CC   which was used for subsequent cassette mutagenesis (along with a vector-
CC   based SacI site) to introduce heterologous C-terminal sequences onto
CC   alphaB-crystallin.
XX
SQ   Sequence 537 BP; 118 A; 171 C; 125 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.22e-53        Length:          537
Score:                    489.00          Matches:         97
Percent Similarity:       73.60%          Conservative:    34
Best Local Similarity:    54.49%          Mismatches:      35
Query Match:              53.38%          Indels:          12
DB:                       10              Gaps:            6

US-10-657-740-1 (1-173) x ADE75374 (1-537)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          9 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 68

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         69 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 125

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||       |||      :::
Db        126 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 185

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 ||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        186 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 245

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        246 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 305

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        306 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 365

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        366 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 425

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                  :::
Db        426 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 473

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||||      |||:::||||||
Db        474 GAGCGCACCATTCCCATCACCCGGGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 527

RESULT 9
ADE75375
ID   ADE75375 standard; DNA; 537 BP.
XX
AC   ADE75375;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human wild-type alphaB-crystallin fragment-encoding DNA.
XX
KW   Hybrid protein chaperone; protein stabilisation; heat shock protein;
KW   sHSP family; protein aggregation inhibition; cell death inhibition;
KW   genome stability pathway inhibition; protein denaturation identification;
KW   protein conformation related disease; cardiomyopathy; cataract;
KW   neurodegenerative disease; cardiant; ophthalmological; neuroprotective;
KW   gene therapy; alphaB-crystallin; wild-type; human; cassette mutagenesis;
KW   ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   mutation        replace(497,G)
FT                   /*tag=  a
FT                   /note= "Mutagenesis of T to G at this site generates a
FT                   unique AvaI site between bases 493-498"
XX
PN   WO2003091266-A2.
XX
PD   06-NOV-2003.
XX
PF   23-APR-2003; 2003WO-GB001721.
XX
PR   23-APR-2002; 2002GB-00009334.
XX
PA   (UYDU-) UNIV DUNDEE.
XX
PI   Quinlan R;
XX
DR   WPI; 2003-865571/80.
XX
PT   New hybrid protein chaperone (e.g. heat shock protein) useful for
PT   stabilizing proteins and/or protein activities, or as an agent to prevent
PT   protein aggregation, or for treating diseases involving altered protein
PT   conformations.
XX
PS   Disclosure; Fig 12; 45pp; English.
XX
CC   The invention relates to a hybrid protein chaperone for stabilising
CC   proteins and/or protein activities. Protein chaperones (also known as
CC   heat shock proteins) are divided into 4 families on the basis of their
CC   primary sequence and chaperone properties: HSP90, HSP70, HSP60 and sHSP
CC   (small heat shock protein). The invention is based upon the finding that
CC   among the sHSP family, which have a general structure of a cenral domain
CC   (called the alpha-crystallin domain) flanked by N and C-terminal regions,
CC   replacement of one or more regions of an sHSP with the corresponding
CC   region from a second sHSP can improve the activity compared to native
CC   sHSPs. In a particualr embodiment of the invention, the hybrid chaperone
CC   is a hybrid sHSP designated alphaB-HSP27 comprising the N-terminus and
CC   central portion of alphaB-crystallin and the C-terminal tail of HSP27.
CC   However, the hybrid protein chaperones of the invention can comprise
CC   regions from HSP90, HSP70 and HSP60 families as well as from the sHSP
CC   family. The invention also encompasses methods for stabilising proteins
CC   such as enzymes, therapuetic proteins, diagnostic proteins, reporter
CC   proteins or antibodies, their fragments or conjugates in an aqueous
```

```
CC   solution using hybrid protein chaperones; stabilised protein formulation
CC   comprising at least one protein associated with the above hybrid protein
CC   chaperone; nucleic acids and vectors encoding a hybrid protein chaperone;
CC   and the coexpression of a recombinant protein of interest and a hybrid
CC   protein chaperone. The hybrid protein chaperones are useful as agents to
CC   prevent protein aggregation, as inhibitors of cell death and genome
CC   stability pathways, for identification of proteins that are in the
CC   process of unfolding, for the treatment of diseases involving altered
CC   protein conformations (e.g., cardiomyopathies, cataracts or
CC   neurodegenerative disease), or for the manufacture of a medicament for
CC   the treatment of such diseases. The present sequence represents a wild-
CC   type human alphaB-crystallin DNA sequence. The T at position 497 was
CC   mutated to G (see ADE75374) to generate a unique AvaI restriction site
CC   which was used for subsequent cassette mutagenesis (along with a vector-
CC   based SacI site) to introduce heterologous C-terminal sequences onto
CC   alphaB-crystallin.
XX
SQ   Sequence 537 BP; 118 A; 171 C; 124 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.22e-53        Length:          537
Score:                    489.00          Matches:         97
Percent Similarity:       73.60%          Conservative:    34
Best Local Similarity:    54.49%          Mismatches:      35
Query Match:              53.38%          Indels:          12
DB:                       10              Gaps:            6

US-10-657-740-1 (1-173) x ADE75375 (1-537)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db          9 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCCTTCTTTCCTTTCCACTCCCCC 68

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         69 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 125

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::       |||       |||       :::
Db        126 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 185

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        186 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 245

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        246 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 305

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::|||||:::|||||||||||||||||||
Db        306 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 365

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        366 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 425

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||||       |||||||:::   |||                 :::
Db        426 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 473

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::||||||||||||||||      |||:::||||||
Db        474 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 527

RESULT 10
AAX39668
ID   AAX39668 standard; DNA; 691 BP.
XX
AC   AAX39668;
XX
DT   02-JUL-1999  (first entry)
XX
DE   Renal cancer associated gene.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US014679.
XX
PR   17-JUL-1997;   97US-00896164.
PR   10-OCT-1997;   97US-0061599P.
PR   10-OCT-1997;   97US-0061765P.
PR   10-OCT-1997;   97US-00948705.
PR   11-OCT-1997;   97GB-00021697.
PR   22-JUN-1998;   98US-00102322.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
PI   O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;
XX
DR   WPI; 1999-132448/11.
XX
PT   New isolated cancer associated nucleic acids and polypeptides - isolated
PT   using sera from cancer patients, used to develop products for the
PT   diagnosis, monitoring or treatment of cancers.
XX
PS   Claim 67; Page 478; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer
XX
SQ   Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.15e-53        Length:          691
Score:                    489.00          Matches:         97
Percent Similarity:       73.60%          Conservative:    34
Best Local Similarity:    54.49%          Mismatches:      35
Query Match:              53.38%          Indels:          12
DB:                       2               Gaps:            6

US-10-657-740-1 (1-173) x AAX39668 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202
```

```
Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db         203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   ||||||||      |||   :::|||
Db         263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db         323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db         383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||      ||||||:::   |||                  :::
Db         443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db         491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 11
ABL65204
ID   ABL65204 standard; DNA; 691 BP.
XX
AC   ABL65204;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:3541.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX
PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0233133P.
PR   18-SEP-2000; 2000US-0233617P.
PR   20-SEP-2000; 2000US-0234009P.
PR   20-SEP-2000; 2000US-0234034P.
PR   20-SEP-2000; 2000US-0234052P.
PR   22-SEP-2000; 2000US-0234509P.
PR   22-SEP-2000; 2000US-0234567P.
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0237172P.
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   03-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 3541; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.15e-53        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x ABL65204 (1-691)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db          26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               |||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db          86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  |||::::::::::||||||:::|||   :::      |||      |||      :::
Db         143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
```

```
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95

Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115

Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135

Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155

Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171

Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 12
ABL62427
ID   ABL62427 standard; DNA; 691 BP.
XX
AC   ABL62427;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Colon adenocarcinoma related gene sequence SEQ ID NO:764.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX
PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0233133P.
PR   18-SEP-2000; 2000US-0233617P.
PR   20-SEP-2000; 2000US-0234009P.
PR   20-SEP-2000; 2000US-0234034P.
PR   20-SEP-2000; 2000US-0234052P.
PR   22-SEP-2000; 2000US-0234509P.
PR   22-SEP-2000; 2000US-0234567P.
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0237172P.
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   03-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 764; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.15e-53        Length:          691
Score:                    489.00          Matches:         97
Percent Similarity:       73.60%          Conservative:    34
Best Local Similarity:    54.49%          Mismatches:      35
Query Match:              53.38%          Indels:          12
DB:                       6               Gaps:            6

US-10-657-740-1 (1-173) x ABL62427 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19

Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39

Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55

Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
```

```
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::|||||::|||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 13
ABT10966
ID   ABT10966 standard; cDNA; 691 BP.
XX
AC   ABT10966;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 1100.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-US002176.
XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS,  Nation M,  Diggans JC,  Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
PS   Claim 1; SEQ ID NO 1100; 260pp + Sequence Listing; English.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-
CC   ABT11112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub.published_pct_sequences
XX
SQ   Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.15e-53        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     6               Gaps:           6

US-10-657-740-1 (1-173) x ABT10966 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              |||::::::::::||||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::|||||::|||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 14
ABK64671
ID   ABK64671 standard; DNA; 691 BP.
XX
AC   ABK64671;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human benign prostatic hyperplasia gene #566.
XX
KW   Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200212440-A2.
XX
PD   14-FEB-2002.
XX
PF   07-AUG-2001; 2001WO-US024708.
XX
PR   07-AUG-2000; 2000US-0223323P.
PR   05-JUN-2001; 2001US-00873319.
XX
```

PA (GENE-) GENE LOGIC INC.
PA (NISB) JAPAN TOBACCO INC.
XX
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX
DR WPI; 2002-257476/30.
XX
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells.
XX
PS Disclosure; Page 307-308; 444pp; English.
XX
CC The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles. (I)
CC is useful for diagnosing the onset or progression of BPH. (II) is useful
CC for identifying an agent that modulates the onset or progression of BPH.
CC The methods are useful to present information identifying the expression
CC level in a tissue or cells, by comparing the expression level of genes
CC given in the specification in the tissue or cells to the level of
CC expression of gene in the database, and displaying the expression levels
CC of at least one gene in the tissue or cell sample compared to the
CC expression level in BPH. Agents using (II) are useful for treating BPH or
CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic
CC hyperplasia gene sequences of the invention
XX
SQ Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.15e-53	Length:	691
Score:	489.00	Matches:	97
Percent Similarity:	73.60%	Conservative:	34
Best Local Similarity:	54.49%	Mismatches:	35
Query Match:	53.38%	Indels:	12
DB:	6	Gaps:	6

US-10-657-740-1 (1-173) x ABK64671 (1-691)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              |||||:::   |||   |||||||||   :::|||         |||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              |||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::|||||          ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||       ||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544
```

RESULT 15
ADF09686
ID ADF09686 standard; cDNA; 691 BP.
XX
AC ADF09686;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human crystallin, alpha B cDNA SEQ ID NO:187.
XX
KW human; protein-protein interaction; virucide; cytostatic; vaccine;
KW human papilloma virus; HPV; cancer; ss; gene.
XX
OS Homo sapiens.
XX
PN WO2003068940-A2.
XX
PD 21-AUG-2003.
XX
PF 14-FEB-2003; 2003WO-US004594.
XX
PR 14-FEB-2002; 2002US-0356911P.
XX
PA (CURA-) CURAGEN CORP.
PA (HOFF) HOFFMANN LA ROCHE INC.
XX
PI Jackson A, Ooi CE, Lewin DA, Cuthill S;
XX
DR WPI; 2003-689668/65.
DR P-PSDB; ADF09576.
XX
PT New purified complex comprising a first polypeptide and a second
PT polypeptide, useful for identifying agents for treating/preventing a
PT condition involving altered level of the complex e.g. human papilloma
PT virus infection, or cancer.
XX
PS Example 3; SEQ ID NO 187; 156pp; English.
XX
CC The invention relates to a novel purified complex comprising a first
CC polypeptide and a second polypeptide, where the polypeptides comprise
CC defined amino acid sequences listed in the specification, and where the
CC first polypeptide binds to the second polypeptide. A complex of the
CC invention has virucide and cytostatic activity, and may have a use as a
CC vaccine. The complex is useful for identifying agents for treating or
CC preventing a conditions involving altered level of the complex, e.g.
CC human papilloma virus (HPV) infection, or cancer. The compositions,
CC antibodies, vectors and methods are useful for treating such diseases.
CC The sequences shown in ADF09584-ADF09697 represent cDNA's of the
CC invention.
XX
SQ Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.15e-53	Length:	691
Score:	489.00	Matches:	97
Percent Similarity:	73.60%	Conservative:	34
Best Local Similarity:	54.49%	Mismatches:	35
Query Match:	53.38%	Indels:	12
DB:	10	Gaps:	6

US-10-657-740-1 (1-173) x ADF09686 (1-691)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||::::::::::||||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::|||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||          ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 16
ADN05800
ID   ADN05800 standard; cDNA; 691 BP.
XX
AC   ADN05800;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic cDNA sequence #1131.
XX
KW   ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   P-PSDB; ADN05801.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 1; SEQ ID NO 2194; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polynucleotides of the invention.
XX
SQ   Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.15e-53      Length:          691
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       12            Gaps:            6

US-10-657-740-1 (1-173) x ADN05800 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||::::::::::||||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::|||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||          ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 17
ADR24825
ID   ADR24825 standard; DNA; 691 BP.
XX
AC   ADR24825;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Breast cancer prognosis marker #686.
XX
KW   ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004065545-A2.
XX
PD   05-AUG-2004.
XX
PF   15-JAN-2004; 2004WO-US001100.
```

```
XX
PR   15-JAN-2003; 2003US-00342887.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
PA   (NECA-) NETHERLANDS CANCER INST.
XX
PI   Van't Veer LJ,  He Y;
XX
DR   WPI; 2004-593473/57.
XX
PT   Classifying a breast cancer patient according to prognosis comprises
PT   determining the similarity between the level of expression of each of
PT   five genes in a cell sample taken from patient, to control levels.
XX
PS   Disclosure; SEQ ID NO 686; 226pp; English.
XX
CC   The invention relates to a method of classifying a breast cancer patient
CC   according to prognosis by determining the similarity between the level of
CC   expression of each of five genes for which markers are listed in the
CC   specification, in a cell sample taken from the breast cancer patient, to
CC   control levels of expression for each respective five genes to obtain a
CC   patient similarity value. The methods are useful for classifying a breast
CC   cancer patient according to prognosis. Kits and computer program products
CC   are useful for data analysis using the diagnostic, prognostic and
CC   statistical methods of the invention. This sequence corresponds to a
CC   marker used in the method of the invention.
XX
SQ   Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.15e-53        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     13              Gaps:           6

US-10-657-740-1 (1-173) x ADR24825 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::|||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   ||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::|||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
```

```
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 18
AAC03893
ID   AAC03893 standard; cDNA; 856 BP.
XX
AC   AAC03893;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 3891.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   P-PSDB; AAG03887.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 3891; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors
XX
SQ   Sequence 856 BP; 201 A; 255 C; 196 G; 197 T; 0 U; 7 Other;

Alignment Scores:
Pred. No.:              8.47e-53        Length:         856
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     3               Gaps:           6

US-10-657-740-1 (1-173) x AAC03893 (1-856)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        229 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 288

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        289 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 345
```

```
Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
               |||:::::::::|||||:::|||   :::      |||      |||      :::
Db         346 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 405

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         406 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 465

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               |||||||||||||||||||||:::|||   ||||||||      |||   :::|||
Db         466 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 525

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db         526 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 585

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db         586 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 645

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||      ||||||:::   |||                 :::
Db         646 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 693

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db         694 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 747

RESULT 19
AAC10867
ID   AAC10867 standard; cDNA; 893 BP.
XX
AC   AAC10867;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14942.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 14942; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 893 BP; 221 A; 259 C; 215 G; 198 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.02e-53        Length:          893
Score:                    489.00          Matches:         97
Percent Similarity:       73.60%          Conservative:    34
Best Local Similarity:    54.49%          Mismatches:      35
Query Match:              53.38%          Indels:          12
DB:                       3               Gaps:            6

US-10-657-740-1 (1-173) x AAC10867 (1-893)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||      .  ||||||:::   |||
Db         266 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 325

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               |||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         326 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 382

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
               |||:::::::::|||||:::|||   :::      |||      |||      :::
Db         383 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 442

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         443 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 502

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               |||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         503 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 562

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db         563 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 622

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db         623 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 682

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||      ||||||:::   |||                 :::
Db         683 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 730

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db         731 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 784

RESULT 20
AAC10863
ID   AAC10863 standard; cDNA; 911 BP.
XX
AC   AAC10863;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14938.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
```

```
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 14938; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 911 BP; 221 A; 260 C; 211 G; 219 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.3e-53       Length:          911
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       3             Gaps:            6

US-10-657-740-1 (1-173) x AAC10863 (1-911)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  |||||||||   :::|||          ||||||:::   |||
Db        284 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 343

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        344 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 400

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        401 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 460

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||:::|||   :::::::|||:::|||   :::   |||
Db        461 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 520

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        521 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 580

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        581 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 640

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        641 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 700

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        701 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 748

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||||      |||:::||||||
Db        749 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 802

RESULT 21
AAC10861
ID   AAC10861 standard; cDNA; 927 BP.
XX
AC   AAC10861;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14936.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 14936; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 927 BP; 230 A; 267 C; 224 G; 206 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.54e-53      Length:          927
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       3             Gaps:            6

US-10-657-740-1 (1-173) x AAC10861 (1-927)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  |||||||||   :::|||          ||||||:::   |||
Db        300 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 359

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              |||||||||||||||||||||||||||||   |||   |||   ||||||   |||
```

```
Db        360 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 416

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        417 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 476

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        477 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 536

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
                 ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        537 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 596

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
                 :::|||||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        597 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 656

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
                 |||:::||||||:::|||:::::||||||         ::::::   |||||||||:::
Db        657 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 716

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
                 ||||||:::||||||      ||||||:::   |||                :::
Db        717 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 764

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
                 ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        765 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 818

RESULT 22
AAC03892
ID   AAC03892 standard; cDNA; 942 BP.
XX
AC   AAC03892;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 3890.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   P-PSDB; AAG03886.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 3890; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors
XX
SQ   Sequence 942 BP; 223 A; 267 C; 224 G; 228 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.::             9.77e-53        Length:         942
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     3               Gaps:           6

US-10-657-740-1 (1-173) x AAC03892 (1-942)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
                ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        315 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 374

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
                |||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        375 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 431

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        432 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 491

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        492 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 551

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
                ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        552 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 611

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
                :::|||||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        612 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 671

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
                 |||:::||||||:::|||:::::||||||         ::::::   |||||||||:::
Db        672 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 731

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
                 ||||||:::||||||      ||||||:::   |||                :::
Db        732 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 779

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
                 ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        780 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 833

RESULT 23
AAZ33574/c
ID   AAZ33574 standard; cDNA; 1036 BP.
XX
AC   AAZ33574;
XX
DT   08-DEC-1999  (first entry)
XX
DE   Human breast tumour-associated EST 34.
XX
KW   Expressed sequence tag; EST; human; breast; cancer; cytostatic;
KW   medicaments; gene therapy; treatment; fat metabolism; ss.
XX
OS   Homo sapiens.
XX
PN   DE19813835-A1.
XX
```

PD 23-SEP-1999.
XX
PF 20-MAR-1998; 98DE-01013835.
XX
PR 20-MAR-1998; 98DE-01013835.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX
DR WPI; 1999-528979/45.
XX
PT Human nucleic acid sequences and protein products from normal breast
PT tissue, useful for breast cancer therapy.
XX
PS Claim 3; 122; 206pp; German.
XX
CC This invention describes novel human nucleic acid sequences frpm normal
CC breast tissue which have cytostatic activity. The nucleic acid sequences
CC can be used to produce and isolate full-length gene sequences. They can
CC be used to express proteins, which can be used as tools to find an
CC activity against breast cancer. The sequences can be used in sense or
CC antisense form. They are especially useful for medicaments for gene
CC therapy to treat breast cancer and for treating illnesses associated with
CC fat metabolism. AAZ33541-Z33610 represent expressed sequence tags
CC described in the method of the invention
XX
SQ Sequence 1036 BP; 240 A; 251 C; 283 G; 262 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.13e-52	Length:	1036
Score:	489.00	Matches:	97
Percent Similarity:	73.60%	Conservative:	34
Best Local Similarity:	54.49%	Mismatches:	35
Query Match:	53.38%	Indels:	12
DB:	2	Gaps:	6

US-10-657-740-1 (1-173) x AAZ33574 (1-1036)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        707 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 648

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        647 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 591

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        590 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 531

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        530 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 471

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        470 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 411

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        410 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 351

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::||||||         ::::::   |||||||||:::
Db        350 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 291

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                 :::
Db        290 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 243

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
```

Db 242 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 189

RESULT 24
ADB52642
ID ADB52642 standard; DNA; 528 BP.
XX
AC ADB52642;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3184.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3184; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 528 BP; 114 A; 170 C; 128 G; 116 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:              1.35e-52        Length:         528
Score:                  485.00          Matches:        96
Percent Similarity:     74.16%          Conservative:   36
Best Local Similarity:  53.93%          Mismatches:     34
Query Match:            52.95%          Indels:         12
DB:                     10              Gaps:           6

US-10-657-740-1 (1-173) x ADB52642 (1-528)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          1 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 ::::::::::::|||||:::|||   :::      |||      |||      :::
Db        118 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        178 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        238 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        298 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 357

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        358 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      |||||||:::   |||            |||:::
Db        418 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 465

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 519
```

```
RESULT 25
ABT41872
ID   ABT41872 standard; DNA; 528 BP.
XX
AC   ABT41872;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Toxicity modelling related rat gene SEQ ID No 1574.
XX
KW   Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW   database; drug screening; toxicity assay; rat; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO200295000-A2.
XX
PD   28-NOV-2002.
XX
PF   22-MAY-2002; 2002WO-US016173.
XX
PR   22-MAY-2001; 2001US-0292335P.
PR   13-JUN-2001; 2001US-0297523P.
PR   19-JUN-2001; 2001US-0298925P.
PR   10-JUL-2001; 2001US-0303807P.
PR   10-JUL-2001; 2001US-0303808P.
PR   10-JUL-2001; 2001US-0303810P.
PR   28-AUG-2001; 2001US-0315047P.
PR   27-SEP-2001; 2001US-0324928P.
PR   22-OCT-2001; 2001US-0330462P.
PR   01-NOV-2001; 2001US-0330867P.
PR   21-NOV-2001; 2001US-0331805P.
PR   06-DEC-2001; 2001US-0336144P.
PR   19-DEC-2001; 2001US-0340873P.
PR   21-FEB-2002; 2002US-0357842P.
PR   21-FEB-2002; 2002US-0357843P.
PR   21-FEB-2002; 2002US-0357844P.
PR   15-MAR-2002; 2002US-0364134P.
PR   08-APR-2002; 2002US-0370144P.
PR   08-APR-2002; 2002US-0370206P.
PR   08-APR-2002; 2002US-0370247P.
PR   17-APR-2002; 2002US-0372794P.
PR   21-APR-2002; 2002US-0371679P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Elashoff M;
XX
DR   WPI; 2003-148464/14.
XX
PT   Predicting at least one toxic effect of a compound, useful for toxicity
PT   modeling, comprises preparing a gene expression profile of a tissue or
PT   cell sample exposed to the compound, and comparing the gene expression
PT   profile to a database.
XX
PS   Example 4; Page; 446pp; English.
XX
CC   The invention relates to a novel method of predicting at least one toxic
CC   effect of a compound. The method comprises a gene expression profile of a
CC   tissue or cell sample exposed to the compound, and comparing the gene
CC   expression profile to a database comprising at least part of the data or
CC   information given in the specification. The methods are useful for
CC   predicting at least one toxic effect of a compound, predicting the
CC   progression of a toxic effect of a compound, predicting the renal
CC   toxicity of a compound, or identifying toxicity markers in tissues or
CC   cells exposed to known renal toxin. The genes are useful as toxicity
CC   markers in drug screening and toxicity assays, in monitoring disease or
CC   physiological states, or disease progression. This polynucleotide
CC   represents a rat DNA sequence relating to the toxic effect database
CC   described in the specification. NOTE: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the World Intellectual Property
CC   Organization
XX
SQ   Sequence 528 BP; 114 A; 170 C; 128 G; 116 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              1.35e-52        Length:         528
Score:                  485.00          Matches:        96
Percent Similarity:     74.16%          Conservative:   36
Best Local Similarity:  53.93%          Mismatches:     34
Query Match:            52.95%          Indels:         12
DB:                     10              Gaps:           6

US-10-657-740-1 (1-173) x ABT41872 (1-528)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          1 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 ::::::::::::|||||:::|||   :::      |||      |||      :::
Db        118 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
```

```
Db        178 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95

Db        238 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115

Db        298 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 357

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135

Db        358 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155

Db        418 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 465

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171

Db        466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 519
```

RESULT 26
ABK63512

```
ID   ABK63512 standard; cDNA; 1247 BP.
XX
AC   ABK63512;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Rat sequence differentially expressed in response to a hepatotoxin #1419.
XX
KW   Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW   differential expression; centrilobular necrosis; steatosis.
XX
OS   Rattus norvegicus.
XX
PN   WO200210453-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US023872.
XX
PR   31-JUL-2000; 2000US-0222040P.
PR   02-NOV-2000; 2000US-0244880P.
PR   11-MAY-2001; 2001US-0290029P.
PR   15-MAY-2001; 2001US-0290645P.
PR   22-MAY-2001; 2001US-0292336P.
PR   06-JUN-2001; 2001US-0295798P.
PR   13-JUN-2001; 2001US-0297457P.
PR   19-JUN-2001; 2001US-0298884P.
PR   09-JUL-2001; 2001US-0303459P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter MW,  Johnson KR,  Castle AL,  Elashoff MR;
XX
DR   WPI; 2002-241625/29.
XX
PT   Predicting toxic effects of compounds or the progression of these toxic
PT   effects by determining the changes in gene expression in tissues or cells
PT   exposed to the toxin and comparing these to gene expression in unexposed
PT   tissues or cells.
XX
PS   Claim 1; SEQ ID NO 1419; 239pp; English.
XX
CC   The invention relates to methods for predicting toxic effects of
CC   compounds or the progression of these toxic effects by determining the
CC   global changes in gene expression in tissues or cells exposed to the
CC   toxin and comparing these to gene expression in unexposed tissues or
CC   cells. Also included are methods of predicting at least one toxic effect
CC   of a compound or progression of a toxic effect, preferably the
CC   hepatotoxicity of a compound, comprising detecting the level of
CC   expression in a tissue or cell sample exposed to the compound of two or
CC   more genes listed in the specification, where differential expression of
CC   the genes is indicative of at least one toxic effect or progression. The
CC   method can also be used to identify an agent which modulates the toxic
CC   response and predict cellular pathways that a compound modulates in a
CC   cell. The methods utilise a set of at least two probes (on a solid
CC   support in kit form), where each of the probes comprises a sequence that
CC   specifically hybridises to a gene listed in the specification, a computer
CC   system comprising a database containing information identifying the
CC   expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC   set of genes comprising at least two genes listed in the specification,
CC   and a user interface to view the information used to present information
CC   identifying the expression level in a tissue or cell of at least one gene
CC   listed in the specification. The method is useful for elucidating global
CC   changes in gene expression and for identifying toxicity markers in
CC   tissues or cell exposed to a known toxin. The genes may be used as
CC   toxicity markers in drug screening and toxicity assays. The genes and
CC   gene expression information may be used as diagnostic markers for the
CC   prediction or identification of the physiological state of tissue or cell
CC   sample that has been exposed to a compound or agent. Hepatotoxicity is
CC   characterised by centrilobular necrosis and steatosis. The present
CC   sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC   which is differentially expressed in response to a hepatotoxic agent
XX
SQ   Sequence 1247 BP; 294 A; 360 C; 306 G; 287 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	4.88e-52	Length:	1247
Score:	485.00	Matches:	96
Percent Similarity:	74.16%	Conservative:	36
Best Local Similarity:	53.93%	Mismatches:	34
Query Match:	52.95%	Indels:	12
DB:	6	Gaps:	6

US-10-657-740-1 (1-173) x ABK63512 (1-1247)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19

Db        583 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 642

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39

Db        643 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 699

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55

Db        700 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 759

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75

Db        760 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 819

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95

Db        820 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 879

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115

Db        880 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 939

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135

Db        940 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 999

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155

Db       1000 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 1047

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171

Db       1048 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 1101
```

```
RESULT 27
ADP72683
ID   ADP72683 standard; DNA; 1247 BP.
XX
AC   ADP72683;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Renal toxin progression gene marker #1272.
XX
KW   ds; toxic effect; gene expression profile; kidney tissue;
KW   differential gene expression; toxicity progression; toxicity marker;
KW   drug screening; toxicity assay; kidney pathology; nephritis;
KW   kidney necrosis; glomerular injury; tubular injury;
KW   focal segmental glomerulosclerosis.
XX
OS   Rattus norvegicus.
XX
PN   WO2004048598-A2.
XX
PD   10-JUN-2004.
XX
PF   24-NOV-2003; 2003WO-US037556.
XX
PR   22-NOV-2002; 2002US-00301856.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick DL,  Porter MW,  Johnson KR,  Castle A,  Higgs B;
PI   Elashoff M;
XX
DR   WPI; 2004-460771/43.
XX
PT   Predicting (the progression of) a toxic effect of a compound, for
PT   monitoring the progression of renal disease states, comprises preparing a
PT   gene expression profile of a kidney tissue or cell sample exposed to the
PT   compound.
XX
PS   Claim 11; SEQ ID NO 1272; 266pp; English.
XX
CC   The invention relates to a method of predicting (the progression of) a
CC   toxic effect of a compound by preparing a gene expression profile of a
CC   kidney tissue or cell sample exposed to the compound and comparing the
CC   gene expression profile to a database, or detecting the level of gene(s)
CC   expression in a tissue or cell sample exposed to the compound, where
CC   differential gene expression compared to a control indicates a toxic
CC   effect (toxicity progression). The method is useful for predicting (the
CC   progression of) at least one toxic effect of a compound. The genes are
CC   useful as toxicity markers in drug screening and toxicity assays. The
CC   methods are useful for predicting the likelihood that a compound or test
CC   agent will induce various specific kidney pathologies, such as nephritis,
CC   kidney necrosis, glomerular and tubular injury, or focal segmental
CC   glomerulosclerosis. The methods are useful for determining the similarity
CC   of a toxic response to one or more individual compounds and for
CC   predicting or elucidating the potential cellular pathways influenced,
CC   induced or modulated by the compound or test agent. The kit is useful for
CC   predicting or modelling the toxic response of a test compound, for
CC   monitoring the progression of renal disease states, for identifying genes
CC   that show promise as new drug targets and for screening known and newly
CC   designed drugs. This sequence corresponds to a gene marker used in the
CC   method of the invention. (Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX
SQ   Sequence 1247 BP; 294 A; 360 C; 306 G; 287 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.88e-52        Length:         1247
Score:                  485.00          Matches:        96
Percent Similarity:     74.16%          Conservative:   36
Best Local Similarity:  53.93%          Mismatches:     34
Query Match:            52.95%          Indels:         12
DB:                     12              Gaps:           6

US-10-657-740-1 (1-173) x ADP72683 (1-1247)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db        583 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 642

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||
Db        643 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 699

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 ::::::::::::|||||:::|||   :::      |||      |||      :::
Db        700 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 759

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        760 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 819

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              |||||||||||||||||||||:::|||   ||||||||      |||   :::|||
Db        820 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 879

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        880 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 939

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::||||||          ::::::   |||||||||:::
Db        940 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 999

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||          |||:::
Db       1000 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 1047

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db       1048 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 1101

RESULT 28
AAC10862
ID   AAC10862 standard; cDNA; 913 BP.
XX
AC   AAC10862;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14937.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST-) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 14937; 71pp + Sequence Listing; English.
```

```
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 913 BP; 228 A; 257 C; 220 G; 202 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.:              1.08e-48      Length:          913
Score:                  457.50        Matches:         94
Percent Similarity:     71.91%        Conservative:    34
Best Local Similarity:  52.81%        Mismatches:      33
Query Match:            49.95%        Indels:          17
DB:                     3             Gaps:            7

US-10-657-740-1 (1-173) x AAC10862 (1-913)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   ||||||:::                   ||||||:::   |||
Db        301 ATGGACATCGCCATCCACCACCCCTTC--------------TTKCCTTTCCACTCCCCC 345

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   |||   |||   |||
Db        346 AGCCGCCTCTTTGACCAGTTCTTCGGRGAGCACCTGTTGGAGTCKGATSTTTTCCCG--- 402

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                |||:::::::::|||||:::|||   :::     |||     |||      :::
Db        403 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 462

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        463 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 522

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        523 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 582

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::|||||:::|||||||||||||||||||
Db        583 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 642

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||         ::::::   |||||||||:::
Db        643 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 702

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                  :::
Db        703 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 750

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db        751 GAGCGCACCATTCCCATCACCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 804

RESULT 29
ABX39112
ID   ABX39112 standard; cDNA; 449 BP.
XX
AC   ABX39112;
XX
DT   20-FEB-2003  (first entry)
XX
DE   Bovine EST associated with lactation/muscle/fat deposition #4277.
XX
KW   Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW   muscle deposition; fat deposition; genome mapping; gene identification;
KW   gene analysis; cattle breeding.
XX
OS   Bos Taurus.
XX
PN   US2002137139-A1.
XX
PD   26-SEP-2002.
XX
PF   24-SEP-2001; 2001US-00960352.
XX
PR   12-JAN-1999;   99US-0115707P.
PR   11-JAN-2000; 2000US-00480902.
XX
PA   (BYAT/) BYATT J C.
PA   (MATH/) MATHIALAGAN N.
PA   (TAON/) TAO N.
PA   (WARR/) WARREN W C.
XX
PI   Byatt JC,  Mathialagan N,  Tao N,  Warren WC;
XX
DR   WPI; 2003-110599/10.
XX
PT   New nucleic acid associated with lactation, and muscle and fat
PT   deposition, useful for genome mapping, gene identification and analysis,
PT   cattle breeding, or for genetically improving cattle.
XX
PS   Claim 2; SEQ ID NO 4277; 245pp; English.
XX
CC   The invention relates to a purified nucleic acid molecule associated with
CC   lactation or muscle and fat deposition (designated LMFD), derived from
CC   cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC   nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC   appearing as ABX34836-ABX49947, or complements of them. Also included are
CC   ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC   acid linked to a promoter and a 3' non- translated sequence that
CC   functions in the cell to cause termination of transcription and addition
CC   of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC   (2) determining a level or pattern of a molecule in a bovine cell or
CC   tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC   of the 15112 nucleic acid sequences or its complement or fragment) with a
CC   complementary nucleic acid molecule obtained from the bovine cell or
CC   tissue, where hybridisation between the marker nucleic acid and the
CC   complementary nucleic acid permits the detection of the molecule; and (b)
CC   detecting the level or pattern of the complementary nucleic acid, where
CC   the detection of the complementary nucleic acid is predictive of the
CC   level or pattern of the molecule. The LMFD nucleic acid is used for
CC   determining a level or pattern of a molecule in a bovine cell or tissue.
CC   It is useful for genome mapping, gene identification and analysis, cattle
CC   breeding, preparation of constructs for use in cattle gene expression, or
CC   for genetically improving cattle. The present sequence is one of the
CC   15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC   present sequence was not shown in the specification but was obtained in
CC   electronic format from the USPTO web site:
CC   seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ   Sequence 449 BP; 91 A; 148 C; 100 G; 110 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.66e-44      Length:          449
Score:                  418.00        Matches:         79
Percent Similarity:     77.30%        Conservative:    30
Best Local Similarity:  56.03%        Mismatches:      26
Query Match:            45.63%        Indels:          6
DB:                     8             Gaps:            4

US-10-657-740-1 (1-173) x ABX39112 (1-449)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         30 ATGGATATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTCCCTTTCCACTCTCCC 89
```

```
Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db          90 AGCCGCCTCTTTGACCAGTTTTTTGGCGAGCACCTGTTAGAGTCTGATCTCTTCCCA--- 146

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  |||::::::::::||||||:::|||   :::      |||      |||      :::
Db         147 GCTTCTACTTCCCTGAGCCCCTTCTACCTTCGGCCGCCCTCATTTCTGCGGGCACCCAGC 206

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                  :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db         207 TGGATTGACACTGGCCTCTCAGAGATGCGTCTGGAGAAGGACAGATTCTCTGTCAACCTG 266

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         267 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAGGTCAAGGTGCTGGGAGATGTGATTGAG 326

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::|||||||||||   |||||||||||:::||||||:::||||||||||||||||||||
Db         327 GTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGGGAGTTCCAC 386

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::|||:::      |||::::::   |||||||||:::
Db         387 AGGAAATACCGGATCCCAGCTGACGTGGAACCTCTCGCCATTACTTCATCCCTGTCGTCT 446

Qy         136 Asp 136
               |||
Db         447 GAT 449
```

RESULT 30
ABX47457
ID ABX47457 standard; cDNA; 396 BP.
XX
AC ABX47457;
XX
DT 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #12622.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 12622; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 396 BP; 80 A; 127 C; 93 G; 96 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.82e-41	Length:	396
Score:	394.00	Matches:	74
Percent Similarity:	77.52%	Conservative:	26
Best Local Similarity:	57.36%	Mismatches:	23
Query Match:	43.01%	Indels:	6
DB:	8	Gaps:	4

US-10-657-740-1 (1-173) x ABX47457 (1-396)

```
Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          13 ATGGATATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTCCCTTTCCACTCTCCC 72

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db          73 AGCCGCCTCTTTGACCAGTTTTTTGGCGAGCACCTGTTGGAGTCTGATCTCTTCCCA--- 129

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  |||::::::::::||||||:::|||   :::      |||      |||      :::
Db         130 GCTTCTACTTCCCTGAGCCCCTTCTACCTTCGGCCGCCCTCATTTCTGCGGGCACCCAGC 189

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                  :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db         190 TGGATTGACACTGGCCTCTCAGAGATGCGTCTGGAGAAGGACAGATTCTCTGTCAACCTG 249

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         250 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAGGTCAAGGTGCTGGGAGATGTGATTGAG 309

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::|||||||||||   |||||||||||:::||||||:::||||||||||||||||||||
Db         310 GTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGGGAGTTCCAC 369

Qy         116 ArgArgTyrArgLeuProSerAsnVal 124
               |||:::||||||:::|||:::::::|||
Db         370 AGGAAATACCGGATCCCAGCTGACGTG 396
```

RESULT 31
ACH93386
ID ACH93386 standard; DNA; 211 BP.
XX
AC ACH93386;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #26581.
XX

```
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 26581; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 211 BP; 32 A; 85 C; 55 G; 39 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.73e-38        Length:          211
Score:                    367.00          Matches:         70
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              40.07%          Indels:          0
DB:                       12              Gaps:            0

US-10-657-740-1 (1-173) x ACH93386 (1-211)

Qy        104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAAC 60

Qy        124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGC 120

Qy        144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGG 180

Qy        164 GluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||
Db        181 GAGGAGAAGCCCACCTCGGCTCCCTCGTCC 210

RESULT 32
ACH79686
ID   ACH79686 standard; DNA; 573 BP.
XX
AC   ACH79686;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #12881.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 12881; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
```

```
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 573 BP; 86 A; 202 C; 171 G; 114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.55e-37      Length:         573
Score:                  367.00        Matches:        70
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            40.07%        Indels:         0
DB:                     12            Gaps:           0

US-10-657-740-1 (1-173) x ACH79686 (1-573)

Qy        104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAAC 262

Qy        124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGC 322

Qy        144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 CCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGG 382

Qy        164 GluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||
Db        383 GAGGAGAAGCCCACCTCGGCTCCCTCGTCC 412

RESULT 33
ACN89210
ID   ACN89210 standard; DNA; 411 BP.
XX
AC   ACN89210;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Breast cancer related marker, seq id 10360.
XX
KW   Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS   Homo sapiens.
XX
PN   US2003099974-A1.
XX
PD   29-MAY-2003.
XX
PF   18-JUL-2002; 2002US-00198846.
XX
PR   18-JUL-2001; 2001US-0306220P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lillie J,  Xu Y,  Wang Y,  Steinmann K;
XX
DR   WPI; 2003-787014/74.
XX
PT   Novel isolated polypeptide associated with breast cancer, useful for
PT   detecting presence of polypeptide in sample, as a marker for breast
PT   cancer.
XX
PS   Disclosure; SEQ ID NO 10360; 36pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) associated with
CC   breast cancer which is encoded by a nucleic acid molecule comprising a
CC   nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC   the polypeptide of the invention. The activity of the polypeptide of the
CC   invention may be described as cytostatic. The antibody is useful for
CC   detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC   invention are useful in the detection of breast tumours. (I) is useful as
CC   a marker for breast cancer and in breast cancer therapy. Sequences given
CC   in records ACN78851-ACN92934 represent nucleic acid markers associated
CC   with breast cancer. Note: The sequence listing does not form part of the
CC   specification but may be obtained in electronic format from the USPTO web
CC   site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ   Sequence 411 BP; 86 A; 134 C; 99 G; 92 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.26e-37      Length:         411
Score:                  364.50        Matches:        71
Percent Similarity:     75.40%        Conservative:   24
Best Local Similarity:  56.35%        Mismatches:     24
Query Match:            39.79%        Indels:         7
DB:                     11            Gaps:           5

US-10-657-740-1 (1-173) x ACN89210 (1-411)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   ||||||||   :::|||        ||||||:::   |||
Db         39 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 98

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||
Db         99 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 155

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              :::::::::::::||||||:::|||   :::      |||      |||      :::
Db        156 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 215

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        216 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 275

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   ||||||||      |||   :::|||
Db        276 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 335

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::|||||:::|||||||||||||||||||
Db        336 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 395

Qy        116 ArgArgTyrArgLeuPro 121
              |||:::|||   ||||||
Db        396 AGGAAGTAC---CTGCCC 410

RESULT 34
ACH17737
ID   ACH17737 standard; cDNA; 450 BP.
XX
AC   ACH17737;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult heart cDNA #2051.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
```

```
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 4949; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 450 BP; 94 A; 144 C; 106 G; 105 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                1.06e-36        Length:          450
Score:                    361.00          Matches:         73
Percent Similarity:       71.94%          Conservative:    27
Best Local Similarity:    52.52%          Mismatches:      33
Query Match:              39.41%          Indels:          6
DB:                       9               Gaps:            4

US-10-657-740-1 (1-173) x ACH17737 (1-450)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||      ||||||   :::|||          |||:::   |||
Db         37 ATGGACATCGCCATCCACCGCCCCTGGATCCGCCGACCCTTCTTTGCTTTCCACTCCCCC 96

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||   |||||||||   |||   |||   ||||||   |||
Db         97 AGCCGCCTCTTTGACCAGTGCTTCGGAGAGCACCTGTCGGAGTCTGATCTTTGCCCG--- 153

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGlnSer------LeuPheArg------Thr 55
                 |||:::::::::||||||:::|||   :::|||       |||   |||       :::
Db        154 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGTCACCCTGCTTACTGCGGGCACCCAGC 213

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||::::::|||   ::::::|||:::      :::   |||

Db        214 TGGTGTGACACTGGACTCTCAAAGATGCGCCTGGAGAAGGACAGGGTCTCTGTCAACCTG 273

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||   |||       |||   :::|||
Db        274 GATGTGAAGCACTTCTCCCCAGAGGAACTCAGAGTTAATGTGTTGTGAGATGAGATTGAG 333

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        334 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 393

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
              |||   ||||||:::|||::::::||||||         ::::::   |||||||||
Db        394 AGGACATACCGGATCCCAGCTGATGTAGACCCTCTGACCATTACTTCATCCCTGTCN 450

RESULT 35
ACH67866
ID   ACH67866 standard; DNA; 579 BP.
XX
AC   ACH67866;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #1061.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 1061; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
```

```
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 579 BP; 89 A; 200 C; 163 G; 127 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              9.84e-35      Length:          579
Score:                  347.00        Matches:         65
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            37.88%        Indels:          0
DB:                     12            Gaps:            0

US-10-657-740-1 (1-173) x ACH67866 (1-579)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 205

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 265

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 325

Qy         61 IleSerGluValArg 65
              |||||||||||||||
Db        326 ATCTCTGAGGTAAGA 340

RESULT 36
ADH76210
ID   ADH76210 standard; DNA; 480 BP.
XX
AC   ADH76210;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human heat shock protein 20 gene SEQ ID NO:303.
XX
KW   heat shock protein 20; cytostatic; antiarteriosclerotic; vasotropic;
KW   antianginal; cerebroprotective; antiarrhythmic; antiasthmatic;
KW   gynaecological; hypotensive; antimigraine; tocolytic; relaxant; HSP;
KW   smooth muscle cell; smooth muscle cell; ds; gene; human.
XX
OS   Homo sapiens.
XX
PN   WO2003018758-A2.
XX
PD   06-MAR-2003.
XX
PF   23-AUG-2002; 2002WO-US026918.
XX
PR   23-AUG-2001; 2001US-0314535P.
XX
PA   (UYAR-) UNIV ARIZONA.
PA   (BROP/) BROPHY C.
PA   (KOMA/) KOMALAVILAS P.
PA   (PANI/) PANITCH A.
PA   (SEAL/) SEAL B.
PA   (LOKE/) LOKESH J.
XX
PI   Brophy C,  Komalavilas P,  Panitch A,  Seal B,  Lokesh J;
XX
DR   WPI; 2003-393248/37.
XX
PT   New heat shock protein 20-derived polypeptides, useful for inhibiting,
PT   treating or preventing smooth muscle cell vasospasm or a disorder such as
PT   intimal hyperplasia, restenosis, atherosclerosis or smooth muscle cell
PT   tumors.
XX
PS   Disclosure; SEQ ID NO 320; 194pp; English.
XX
CC   The invention relates to a novel polypeptide comprising a heat shock
CC   protein 20-derived polypeptide. A polypeptide of the invention has
CC   cytostatic, antiarteriosclerotic, vasotropic, antianginal,
CC   cerebroprotective, antiarrhythmic, antiasthmatic, gynaecological,
CC   hypotensive, antimigraine, tocolytic, and relaxant activity, and may act
CC   as a HSP agonist or antagonist. The polypeptides, heat shock protein
CC   (HSP) 20, and methods are useful for treating or preventing a disorder,
CC   e.g. intimal hyperplasia, stenosis, restenosis, transplant vasculopathy,
CC   atherosclerosis, smooth muscle cell tumours such as leiomyosarcoma, or
CC   vasospasm, which is associated with angina, coronary vasospasm,
CC   Prinzmetal's angina, coronary ischaemia, stroke, bradycardia,
CC   hypertension, pulmonary (lung) hypertension, asthma (bronchospasm),
CC   toxemia of pregnancy, pre-term labour, pre-eclampsia/eclampsia, Raynaud's
CC   disease, Raynaud's phenomenon, haemolytic-uremia, non-occlusive
CC   mesenteric ischaemia, anal fissure, achalasia, impotence, migraine, or
CC   ischaemic muscle injury associated with smooth muscle spasm. The
CC   polypeptides and HSP 20 are also useful for inhibiting smooth muscle cell
CC   proliferation and/or migration. The present sequence represents a peptide
CC   used in a polypeptide of the invention.
XX
SQ   Sequence 480 BP; 110 A; 133 C; 131 G; 106 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.95e-34      Length:          480
Score:                  340.00        Matches:         69
Percent Similarity:     58.14%        Conservative:    31
Best Local Similarity:  40.12%        Mismatches:      50
Query Match:            37.12%        Indels:          22
DB:                     10            Gaps:            3

US-10-657-740-1 (1-173) x ADH76210 (1-480)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
              |||   :::|||       |||   :::|||         |||
Db         13 GTTCCAGTCCAGCCTAGTTGGCTAAGAAGAGCTAGTGCGCCTTTGCCGGGTTTGAGTGCC 72

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   |||||||||||||||   ||||||||||||||   |||  :::|||
Db         73 CCCGGGAGGCTATTTGATCAACGCTTTGGCGAGGGGTTACTGGAGGCTGAATTAGCAGCA 132

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                     :::|||::::::|||||||||         |||           :::
Db        133 CTTTGTCCGACCACACTCGCGCCCTATTAC---------CTTAGAGCGCCGTCTGTAGCC 183

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                    :::::::::|||   :::|||          |||   :::   ||||||||||||
Db        184 TTACCAGTCGCTCAGGTACCAACTGACCCAGGCCACTTCTCCGTTTTATTAGACGTGAAA 243

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        244 CACTTTAGCCCAGAAGAGATAGCAGTCAAAGTTGTAGGAGAGCATGTGGAAGTTCACGCG 303

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||||
Db        304 AGACATGAAGAGAGACCAGATGAACATGGTTTCGTAGCGAGAGAATTCCATCGACGGTAT 363

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              |||||||||      ||||||   :::|||::::::   :::|||||||   :::|||:::
Db        364 CGTCTGCCCCCAGGAGTCGATCCTGCAGCTGTGACGAGTGCATTATCGCCTGAGGGAGTG 423
```

```
Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                 :::|||
Db        424 CTCAGTATC-------------------------------------------CAAGCA 438

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                 |||   |||   :::      |||   :::|||
Db        439 GCCCCCGCGTCAGCCCAAGCCCCGCCTCCGGCTGCT 474
```

RESULT 37
ABQ61117
ID ABQ61117 standard; cDNA; 575 BP.
XX
AC ABQ61117;
XX
DT 26-FEB-2003 (first entry)
XX
DE Skeletal muscle stress protein p20 encoding sequence.
XX
KW Neuroprotective; immunomodulator; cancer; chromosome 19pter-q12;
KW cytostatic; anti-inflammatory; gene therapy; nutritional supplement;
KW wound; burn; ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnerary; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231111-A2.
XX
PD 18-APR-2002.
XX
PF 11-OCT-2001; 2001WO-US027760.
XX
PR 12-OCT-2000; 2000US-00687527.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-426278/45.
DR N-PSDB; ABP43873.
XX
PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
PS Claim 1; SEQ ID # 330; 357pp + Sequence Listing; English.
XX
CC The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnerary, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records ABQ60788-
CC ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 575 BP; 87 A; 216 C; 183 G; 89 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.79e-34	Length:	575
Score:	340.00	Matches:	69
Percent Similarity:	58.14%	Conservative:	31
Best Local Similarity:	40.12%	Mismatches:	50
Query Match:	37.12%	Indels:	22
DB:	6	Gaps:	3

US-10-657-740-1 (1-173) x ABQ61117 (1-575)

```
Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||       |||   :::|||       |||
Db         16 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 75

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   |||||||||||||||   ||||||||||||||||   |||   :::|||
Db         76 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 135

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                  :::|||::::::|||||||||          |||            :::
Db        136 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 186

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                 :::::::::|||   :::|||       |||   :::   ||||||||||
Db        187 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 246

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||:::::   |||||||||       :::   ||||||:::|||
Db        247 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 306

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||
Db        307 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 366

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      ||||||   :::|||::::::   :::|||||||   :::|||:::
Db        367 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 426

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                 :::|||
Db        427 CTGTCCATC-------------------------------------------CAGGCC 441

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                 |||   |||   :::      |||   :::|||
Db        442 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 477
```

RESULT 38
ADM19230
ID ADM19230 standard; cDNA; 1433 BP.
XX
AC ADM19230;
XX
DT 20-MAY-2004 (first entry)
XX
DE Novel human channel/transporter gene #27.
XX
KW ds; gene; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW opthalmological; gene therapy; channel/transporter protein;
KW rheumatoid arthritis; neoplasm; cardiac arrest; cerebrovascular disorder;
KW cerebral ischemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn; transplantation;
KW chemotaxis; food additive.
XX
OS Homo sapiens.
XX
PN WO200154472-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001307.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476159/51.
DR P-PSDB; ADM19709.
XX
PT Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
PS Claim 1; SEQ ID NO 37; 809pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to

```
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. The antibodies to the proteins can also be used
CC   in alleviating symptoms associated with the disorders and in diagnostic
CC   immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC   (ELISA). Disorders which are diagnosed or treated include autoimmune
CC   diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC   neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC   arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
CC   nervous system disorders e.g. Alzheimer's disease, infections caused by
CC   bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC   The polypeptides can also be used to aid wound healing and epithelial
CC   cell proliferation, to prevent skin aging due to sunburn, to maintain
CC   organs before transplantation, for supporting cell culture of primary
CC   tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC   also be used as a food additive or preservative to increase or decrease
CC   storage capabilities. This sequence corresponds to a gene of the
CC   invention.
XX
SQ   Sequence 1433 BP; 260 A; 592 C; 307 G; 263 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:                3.05e-33      Length:          1433
Score:                    340.00        Matches:         69
Percent Similarity:       58.14%        Conservative:    31
Best Local Similarity:    40.12%        Mismatches:      50
Query Match:              37.12%        Indels:          22
DB:                       5             Gaps:            3

US-10-657-740-1 (1-173) x ADM19230 (1-1433)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||      |||   :::|||         |||
Db         43 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 102

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db        103 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 162

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                    :::|||:::::|||||||||         |||            :::
Db        163 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 213

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  :::::::::|||   :::|||         |||   :::   ||||||||||||
Db        214 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 273

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||:::::   |||||||||      :::   ||||||:::|||
Db        274 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 333

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||||
Db        334 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 393

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      ||||||   :::|||::::::   :::||||||   :::|||:::
Db        394 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 453

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                :::|||
Db        454 CTGTCCATC-------------------------------------------CAGGCC 468

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                |||   |||   :::      |||   :::|||
Db        469 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 504

RESULT 39
AAZ42230
ID   AAZ42230 standard; cDNA; 1488 BP.
XX
AC   AAZ42230;
XX
DT   31-JAN-2000  (first entry)
XX
DE   Human normal bladder tissue cDNA derived EST 109.
XX
KW   Human; bladder; treatment; EST; expressed sequence tag; cytostatic;
KW   cancer; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   DE19818620-A1.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1998;   98DE-01018620.
XX
PR   21-APR-1998;   98DE-01018620.
XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Rosenthal A,  Specht T,  Hinzmann B,  Schmitt A,  Pilarsky C,  Dahl E;
XX
DR   WPI; 1999-602416/52.
DR   P-PSDB; AAY60544, AAY60545, AAY60546.
XX
PT   New polypeptides and their nucleic acids, useful for treatment of bladder
PT   tumor and identification of therapeutic agents.
XX
PS   Claim 3; Page 230; 366pp; German.
XX
CC   This invention describes novel polypeptide fragment sequences (I) and
CC   their encoding nucleic acids (II) which are highly expressed in normal
CC   bladder tissue and have cytostatic activity. (II) are used for
CC   recombinant expression of (I) and to isolate complete genes. (I) are used
CC   to identify agents suitable for the treatment of bladder tumours, to
CC   directly treat this form of cancer (including expression from gene
CC   therapy vectors), or are used in a preparation for cancer treatment. (I)
CC   is also used for the generation of specific antibodies. (II) are
CC   identified by assembling ESTs (expressed sequence tags) from a particular
CC   tissue type before comparison of expression patterns. This allows a
CC   significantly longer fragment of the gene to be revealed, and therefore
CC   reduces the number of failures because of ESTs from different libraries
CC   representing different parts of the same unknown gene distorting the
CC   estimated frequency of occurrence in a particular tissue. AAZ42122-Z42248
CC   represent EST fragments derived from a human normal bladder tissue cDNA
CC   library which encode the protein fragments represented in AAY60329-Y60591
XX
SQ   Sequence 1488 BP; 287 A; 607 C; 321 G; 273 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.23e-33      Length:          1488
Score:                    340.00        Matches:         69
Percent Similarity:       58.14%        Conservative:    31
Best Local Similarity:    40.12%        Mismatches:      50
Query Match:              37.12%        Indels:          22
DB:                       2             Gaps:            3

US-10-657-740-1 (1-173) x AAZ42230 (1-1488)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||      |||   :::|||         |||
Db         29 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 88

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db         89 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 148

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                    :::|||:::::|||||||||         |||            :::
Db        149 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 199

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  :::::::::|||   :::|||         |||   :::   ||||||||||||
```

```
Db        200 CTGCCCGTGGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 259

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        260 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 319

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||
Db        320 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 379

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      |||||   :::|||::::::   :::||||||   :::|||:::
Db        380 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 439

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                :::|||
Db        440 CTGTCCATC----------------------------------------------CAGGCC 454

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                |||   |||   :::      |||   :::|||
Db        455 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 490

RESULT 40
ADM19501/c
ID   ADM19501 standard; cDNA; 1506 BP.
XX
AC   ADM19501;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Novel human channel/transporter gene #27 clone 2.
XX
KW   ds; gene; immunosuppressive; antiarthritic; antirheumatic;
KW   antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW   nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW   opthalmological; gene therapy; channel/transporter protein;
KW   rheumatoid arthritis; neoplasm; cardiac arrest; cerebrovascular disorder;
KW   cerebral ischemia; angiogenesis; nervous system disorder;
KW   Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KW   epithelial cell proliferation; skin aging; sunburn; transplantation;
KW   chemotaxis; food additive.
XX
OS   Homo sapiens.
XX
PN   WO200154472-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001307.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
```

```
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-476159/51.
DR   P-PSDB; ADM19980.
XX
PT   Isolated nucleic acid molecule encoding a channel/transporter protein is
PT   used in preventing, treating or ameliorating a medical condition.
XX
PS   Claim 1; SEQ ID NO 308; 809pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encoding a
CC   channel/transporter protein or sequences at least 95% identical to a
CC   these. The nucleic acids and proteins encoded by them are used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. The antibodies to the proteins can also be used
CC   in alleviating symptoms associated with the disorders and in diagnostic
CC   immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC   (ELISA). Disorders which are diagnosed or treated include autoimmune
CC   diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC   neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC   arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
CC   nervous system disorders e.g. Alzheimer's disease, infections caused by
CC   bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC   The polypeptides can also be used to aid wound healing and epithelial
CC   cell proliferation, to prevent skin aging due to sunburn, to maintain
CC   organs before transplantation, for supporting cell culture of primary
CC   tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC   also be used as a food additive or preservative to increase or decrease
CC   storage capabilities. This sequence corresponds to a gene of the
CC   invention.
XX
SQ   Sequence 1506 BP; 276 A; 318 C; 616 G; 295 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              3.28e-33        Length:         1506
Score:                  340.00          Matches:        69
Percent Similarity:     58.14%          Conservative:   31
Best Local Similarity:  40.12%          Mismatches:     50
Query Match:            37.12%          Indels:         22
DB:                     5               Gaps:           3

US-10-657-740-1 (1-173) x ADM19501 (1-1506)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||     |||   :::|||       |||
Db       1475 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 1416

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||||   ||||||||||||||||   |||   :::|||
Db       1415 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 1356

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                    :::|||::::::||||||||          |||           :::
Db       1355 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 1305

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  :::::::::|||   :::|||       |||   :::   ||||||||||||
Db       1304 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 1245

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||       :::   ||||||:::|||
Db       1244 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 1185

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||||
Db       1184 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 1125

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      ||||||  :::|||::::::   :::|||||   :::|||:::
Db       1124 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 1065

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                :::|||
Db       1064 CTGTCCATC-----------------------------------------------CAGGCC 1050

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                |||   |||   :::      |||   :::|||
Db       1049 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 1014

RESULT 41
AAC76681
ID   AAC76681 standard; cDNA; 1771 BP.
XX
AC   AAC76681;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2236 polynucleotide sequence SEQ ID NO:4471.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
```

```
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US008621.
XX
PR   31-MAR-1999;   99US-0127607P.
PR   02-APR-1999;   99US-0127636P.
PR   05-APR-1999;   99US-0127728P.
PR   30-MAR-2000; 2000US-00540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB42472.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease.
XX
PS   Claim 5; Page 3656-3657; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC   anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC   cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC   dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC   antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC   sequences can be used for determining the presence of or predisposition
CC   to, or preventing or treating pathological conditions associated with an
CC   ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC   proteins in gene therapy vectors. The proteins and nucleic acids may be
CC   used to treat cancers, proliferative disorders, neurodegenerative
CC   disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC   diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC   storage, systemic lupus erythematosus, severe combined immunodeficiency
CC   (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC   disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC   cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC   enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ   Sequence 1771 BP; 342 A; 704 C; 389 G; 336 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.18e-33        Length:         1771
Score:                  340.00          Matches:        69
Percent Similarity:     58.14%          Conservative:   31
Best Local Similarity:  40.12%          Mismatches:     50
Query Match:            37.12%          Indels:         22
DB:                     3               Gaps:           3

US-10-657-740-1 (1-173) x AAC76681 (1-1771)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
              |||   :::|||      |||   :::|||         |||
Db        313 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 372

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db        373 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 432

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                 :::|||::::::|||||||||         |||             :::
Db        433 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 483

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                 :::::::::|||   :::|||         |||   :::   ||||||||||
Db        484 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 543

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        544 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 603

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::||||||||||||||||||||
Db        604 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 663

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      ||||||   :::|||::::::   :::||||||   :::|||:::
Db        664 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 723

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                :::|||
Db        724 CTGTCCATC----------------------------------------------CAGGCC 738

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                |||   |||   :::      |||   :::|||
Db        739 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 774

RESULT 42
ADQ22255
ID   ADQ22255 standard; DNA; 1820 BP.
XX
AC   ADQ22255;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 5075.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 5075; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
```

```
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 1820 BP; 374 A; 822 C; 335 G; 289 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.36e-33        Length:         1820
Score:                  340.00          Matches:        69
Percent Similarity:     58.14%          Conservative:   31
Best Local Similarity:  40.12%          Mismatches:     50
Query Match:            37.12%          Indels:         22
DB:                     12              Gaps:           3

US-10-657-740-1 (1-173) x ADQ22255 (1-1820)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||       |||   :::|||         |||
Db         35 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 94

Qy         19 ProSerArgLeuPheAspGlnPheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   |||||||||||||||   ||||||||||||||   |||   :::|||
Db         95 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 154

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                    :::|||::::::|||||||||         |||           :::
Db        155 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 205

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  :::::::::|||   :::|||         |||   :::   ||||||||||
Db        206 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 265

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        266 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 325

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||::::::::::||||||||||||||||||||
Db        326 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 385

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              |||||||||       ||||||   :::|||::::::   :::||||||   :::|||:::
Db        386 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 445

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                :::|||
Db        446 CTGTCCATC-------------------------------------------CAGGCC 460

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
                |||   |||   :::       |||   :::|||
Db        461 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 496

RESULT 43
ACH81569
ID   ACH81569 standard; DNA; 212 BP.
XX
AC   ACH81569;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #14764.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 14764; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 212 BP; 33 A; 81 C; 49 G; 49 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.18e-34        Length:         212
Score:                  338.00          Matches:        63
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            36.90%          Indels:         0
DB:                     12              Gaps:           0

US-10-657-740-1 (1-173) x ACH81569 (1-212)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 83

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 143

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        144 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 203

Qy         61 IleSerGlu 63
              |||||||||
Db        204 ATCTCTGAG 212

RESULT 44
ACH17560
ID   ACH17560 standard; cDNA; 421 BP.
XX
AC   ACH17560;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult heart cDNA #1874.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 4772; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 421 BP; 90 A; 133 C; 94 G; 104 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.86e-34        Length:         421
Score:                  338.00          Matches:        68
Percent Similarity:     72.26%          Conservative:   31
Best Local Similarity:  49.64%          Mismatches:     32
Query Match:            36.90%          Indels:         7
DB:                     9               Gaps:           4

US-10-657-740-1 (1-173) x ACH17560 (1-421)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrPro--- 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::
Db         15 ATGGACATCGCCATCCACCACCCTTGGATCCGCCGGCCCTTCTTTCCTTTTCACTCCCGC 74

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||   |||   ||||||||||||   |||   |||   ||||||   |||
Db         75 GGCCGCCTCTTAGACCATTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 131

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        132 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTTCTGCGGGCACCCAGC 191

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::|||::::::|||   ::::::||||||      :::   |||
Db        192 TGGGTTGACACTGGACTCTCAAAGATGCGCCTGGAGAAGGACAAGATCTCTGTCAACCTG 251

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||:::::|||      |||||      |||   ::::::
Db        252 GATGTGAAGCACTTCTCCCCAAAGGAACTCAAAGGTAAGGTGTG-GGAGACGTGATTGAC 310

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||:::|||||||||||:::||||||:::||||||||||||||||||
Db        311 GTGCATGGTAAACATGACGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 370

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 132
              |||:::||||||:::|||::::::||||||         ::::::   |||
Db        371 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTTACCATTACTTCATCC 421

RESULT 45
ADH76202
ID   ADH76202 standard; DNA; 486 BP.
XX
AC   ADH76202;
XX
DT   22-APR-2004  (first entry)
XX
DE   Rat heat shock protein 20 gene SEQ ID NO:303.
XX
KW   heat shock protein 20; cytostatic; antiarteriosclerotic; vasotropic;
KW   antianginal; cerebroprotective; antiarrhythmic; antiasthmatic;
KW   gynaecological; hypotensive; antimigraine; tocolytic; relaxant; HSP;
KW   smooth muscle cell; smooth muscle cell; ds; gene.
XX
OS   Rattus sp.
XX
PN   WO2003018758-A2.
XX
PD   06-MAR-2003.
XX
PF   23-AUG-2002; 2002WO-US026918.
XX
PR   23-AUG-2001; 2001US-0314535P.
XX
PA   (UYAR-) UNIV ARIZONA.
PA   (BROP/) BROPHY C.
PA   (KOMA/) KOMALAVILAS P.
PA   (PANI/) PANITCH A.
PA   (SEAL/) SEAL B.
PA   (LOKE/) LOKESH J.
XX
PI   Brophy C,  Komalavilas P,  Panitch A,  Seal B,  Lokesh J;
XX
DR   WPI; 2003-393248/37.
XX
PT   New heat shock protein 20-derived polypeptides, useful for inhibiting,
PT   treating or preventing smooth muscle cell vasospasm or a disorder such as
PT   intimal hyperplasia, restenosis, atherosclerosis or smooth muscle cell
PT   tumors.
XX
PS   Disclosure; SEQ ID NO 303; 194pp; English.
```

```
XX
CC   The invention relates to a novel polypeptide comprising a heat shock
CC   protein 20-derived polypeptide. A polypeptide of the invention has
CC   cytostatic, antiarteriosclerotic, vasotropic, antianginal,
CC   cerebroprotective, antiarrhythmic, antiasthmatic, gynaecological,
CC   hypotensive, antimigraine, tocolytic, and relaxant activity, and may act
CC   as a HSP agonist or antagonist. The polypeptides, heat shock protein
CC   (HSP) 20, and methods are useful for treating or preventing a disorder,
CC   e.g. intimal hyperplasia, stenosis, restenosis, transplant vasculopathy,
CC   atherosclerosis, smooth muscle cell tumours such as leiomyosarcoma, or
CC   vasospasm, which is associated with angina, coronary vasospasm,
CC   Prinzmetal's angina, coronary ischaemia, stroke, bradycardia,
CC   hypertension, pulmonary (lung) hypertension, asthma (bronchospasm),
CC   toxemia of pregnancy, pre-term labour, pre-eclampsia/eclampsia, Raynaud's
CC   disease, Raynaud's phenomenon, haemolytic-uremia, non-occlusive
CC   mesenteric ischaemia, anal fissure, achalasia, impotence, migraine, or
CC   ischaemic muscle injury associated with smooth muscle spasm. The
CC   polypeptides and HSP 20 are also useful for inhibiting smooth muscle cell
CC   proliferation and/or migration. The present sequence represents a peptide
CC   used in a polypeptide of the invention.
XX
SQ   Sequence 486 BP; 109 A; 139 C; 128 G; 110 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.99e-33        Length:         486
Score:                  336.00          Matches:        71
Percent Similarity:     57.14%          Conservative:   29
Best Local Similarity:  40.57%          Mismatches:     53
Query Match:            36.68%          Indels:         22
DB:                     10              Gaps:           3

US-10-657-740-1 (1-173) x ADH76202 (1-486)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||      |||   :::|||          |||
Db         13 GTGCCGGTACAACCCAGCTGGCTGCGGCGTGCTTCCGCGCCATTACCTGGCTTCAGTACC 72

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   |||||||||||||||   ||||||||||||||||   |||   :::|||
Db         73 CCCGGACGATTGTTTGACCAGAGGTTTGGGGAAGGTTTACTTGAGGCGGAATTAGCAAGT 132

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                  :::   |||:::||||||||      |||          :::
Db        133 CTATGTCCTGCAGCTATAGCACCCTACTAC---------CTAAGGGCACCATCTGTCGCG 183

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  ::::::|||   :::|||          |||   :::   ||||||||||
Db        184 CTCCCAACTGCCCAAGTGCCCACGGATCCAGGCTATTTCAGCGTTCTGTTAGACGTAAAG 243

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              |||||||||||||||:::::::::||||||||       |||   ||||||:::|||
Db        244 CATTTTAGTCCAGAAGAAATTTCAGTAAAAGTAGTGGGAGACCATGTCGAGGTACATGCT 303

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::|||:::||||||||||||||||||||
Db        304 AGACACGAAGAGAGACCTGATGAACACGGTTTCATCGCTCGAGAGTTTCACCGGCGTTAT 363

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      ||||||   :::|||::::::   :::||||||   :::|||:::
Db        364 CGCTTGCCGCCGGGGGTTGATCCCGCGGCCGTCACATCAGCACTCAGTCCGGAGGGAGTT 423

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                              :::|||
Db        424 TTATCCATA-----------------------------------------------CAAGCC 438

Qy        159 IleProValSerArgGluGluLysProThrSerAlaProSerSer 173
                 |||   |||   :::           |||   |||::::::
Db        439 ACACCGGCCTCTGCTCAGGCCTCGCTTCCATCGCCTCCTGCGGCA 483

RESULT 46
ABT41710
ID   ABT41710 standard; DNA; 1310 BP.
XX
AC   ABT41710;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Toxicity modelling related rat gene SEQ ID No 1412.
XX
KW   Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW   database; drug screening; toxicity assay; rat; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO200295000-A2.
XX
PD   28-NOV-2002.
XX
PF   22-MAY-2002; 2002WO-US016173.
XX
PR   22-MAY-2001; 2001US-0292335P.
PR   13-JUN-2001; 2001US-0297523P.
PR   19-JUN-2001; 2001US-0298925P.
PR   10-JUL-2001; 2001US-0303807P.
PR   10-JUL-2001; 2001US-0303808P.
PR   10-JUL-2001; 2001US-0303810P.
PR   28-AUG-2001; 2001US-0315047P.
PR   27-SEP-2001; 2001US-0324928P.
PR   22-OCT-2001; 2001US-0330462P.
PR   01-NOV-2001; 2001US-0330867P.
PR   21-NOV-2001; 2001US-0331805P.
PR   06-DEC-2001; 2001US-0336144P.
PR   19-DEC-2001; 2001US-0340873P.
PR   21-FEB-2002; 2002US-0357842P.
PR   21-FEB-2002; 2002US-0357843P.
PR   21-FEB-2002; 2002US-0357844P.
PR   15-MAR-2002; 2002US-0364134P.
PR   08-APR-2002; 2002US-0370144P.
PR   08-APR-2002; 2002US-0370206P.
PR   08-APR-2002; 2002US-0370247P.
PR   17-APR-2002; 2002US-0372794P.
PR   21-APR-2002; 2002US-0371679P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Elashoff M;
XX
DR   WPI; 2003-148464/14.
XX
PT   Predicting at least one toxic effect of a compound, useful for toxicity
PT   modeling, comprises preparing a gene expression profile of a tissue or
PT   cell sample exposed to the compound, and comparing the gene expression
PT   profile to a database.
XX
PS   Example 4; Page; 446pp; English.
XX
CC   The invention relates to a novel method of predicting at least one toxic
CC   effect of a compound. The method comprises a gene expression profile of a
CC   tissue or cell sample exposed to the compound, and comparing the gene
CC   expression profile to a database comprising at least part of the data or
CC   information given in the specification. The methods are useful for
CC   predicting at least one toxic effect of a compound, predicting the
CC   progression of a toxic effect of a compound, predicting the renal
CC   toxicity of a compound, or identifying toxicity markers in tissues or
CC   cells exposed to known renal toxin. The genes are useful as toxicity
CC   markers in drug screening and toxicity assays, in monitoring disease or
CC   physiological states, or disease progression. This polynucleotide
CC   represents a rat DNA sequence relating to the toxic effect database
CC   described in the specification. NOTE: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the World Intellectual Property
CC   Organization
XX
SQ   Sequence 1310 BP; 274 A; 446 C; 280 G; 310 T; 0 U; 0 Other;

Alignment Scores:
```

```
Pred. No.:              8.75e-33        Length:         1310
Score:                  336.00          Matches:        71
Percent Similarity:     57.14%          Conservative:   29
Best Local Similarity:  40.57%          Mismatches:     53
Query Match:            36.68%          Indels:         22
DB:                     10              Gaps:           3

US-10-657-740-1 (1-173) x ABT41710 (1-1310)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
              |||   :::|||      |||   :::|||         |||
Db         18 GTGCCTGTGCAGCCTTCTTGGCTGCGCCGTGCTTCAGCTCCTTTACCGGGTTTTTCCACT 77

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db         78 CCGGGACGCCTCTTTGACCAGCGTTTCGGCGAAGGGCTGCTTGAGGCAGAGCTGGCTTCA 137

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                     :::   |||:::|||||||||         |||             :::
Db        138 CTGTGCCCTGCTGCGATCGCGCCCTACTAT---------CTGCGCGCCCCCAGTGTGGCG 188

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                     :::::::|||   :::|||         |||   :::   ||||||||||||
Db        189 TTACCCACAGCCCAGGTGCCCACGGACCCTGGGTATTTTTCTGTGCTGCTGGATGTGAAG 248

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              |||||||||||||:::::::::|||||||||       |||   ||||||:::|||
Db        249 CACTTCTCGCCAGAGGAAATCTCTGTCAAGGTGGTTGGTGACCATGTGGAGGTCCATGCT 308

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::|||:::|||||||||||||||||||||
Db        309 CGGCATGAGGAGCGCCCAGATGAACATGGATTCATTGCTCGAGAGTTCCACCGCCGATAC 368

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              |||||||||       ||||||   :::|||::::::   :::||||||   :::|||:::
Db        369 CGCCTGCCTCCTGGCGTGGACCCTGCTGCAGTGACCTCTGCACTGTCTCCTGAGGGTGTT 428

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                              :::|||
Db        429 CTGTCTATC-----------------------------------------------CAGGCC 443

Qy        159 IleProValSerArgGluGluLysProThrSerAlaProSerSer 173
                 |||   |||   :::            |||   |||::::::
Db        444 ACACCAGCGTCGGCCCAGGCCTCACTTCCGTCACCACCTGCTGCC 488

RESULT 47
ABX38978
ID   ABX38978 standard; cDNA; 380 BP.
XX
AC   ABX38978;
XX
DT   20-FEB-2003  (first entry)
XX
DE   Bovine EST associated with lactation/muscle/fat deposition #4143.
XX
KW   Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW   muscle deposition; fat deposition; genome mapping; gene identification;
KW   gene analysis; cattle breeding.
XX
OS   Bos Taurus.
XX
PN   US2002137139-A1.
XX
PD   26-SEP-2002.
XX
PF   24-SEP-2001; 2001US-00960352.
XX
PR   12-JAN-1999;   99US-0115707P.
PR   11-JAN-2000; 2000US-00480902.
XX
PA   (BYAT/) BYATT J C.
PA   (MATH/) MATHIALAGAN N.
PA   (TAON/) TAO N.
PA   (WARR/) WARREN W C.
XX
PI   Byatt JC,  Mathialagan N,  Tao N,  Warren WC;
XX
DR   WPI; 2003-110599/10.
XX
PT   New nucleic acid associated with lactation, and muscle and fat
PT   deposition, useful for genome mapping, gene identification and analysis,
PT   cattle breeding, or for genetically improving cattle.
XX
PS   Claim 2; SEQ ID NO 4143; 245pp; English.
XX
CC   The invention relates to a purified nucleic acid molecule associated with
CC   lactation or muscle and fat deposition (designated LMFD), derived from
CC   cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC   nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC   appearing as ABX34836-ABX49947, or complements of them. Also included are
CC   ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC   acid linked to a promoter and a 3' non- translated sequence that
CC   functions in the cell to cause termination of transcription and addition
CC   of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC   (2) determining a level or pattern of a molecule in a bovine cell or
CC   tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC   of the 15112 nucleic acid sequences or its complement or fragment) with a
CC   complementary nucleic acid molecule obtained from the bovine cell or
CC   tissue, where hybridisation between the marker nucleic acid and the
CC   complementary nucleic acid permits the detection of the molecule; and (b)
CC   detecting the level or pattern of the complementary nucleic acid, where
CC   the detection of the complementary nucleic acid is predictive of the
CC   level or pattern of the molecule. The LMFD nucleic acid is used for
CC   determining a level or pattern of a molecule in a bovine cell or tissue.
CC   It is useful for genome mapping, gene identification and analysis, cattle
CC   breeding, preparation of constructs for use in cattle gene expression, or
CC   for genetically improving cattle. The present sequence is one of the
CC   15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC   present sequence was not shown in the specification but was obtained in
CC   electronic format from the USPTO web site:
CC   seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ   Sequence 380 BP; 90 A; 114 C; 101 G; 75 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.89e-33        Length:         380
Score:                  332.50          Matches:        65
Percent Similarity:     75.65%          Conservative:   22
Best Local Similarity:  56.52%          Mismatches:     23
Query Match:            36.30%          Indels:         5
DB:                     8               Gaps:           2

US-10-657-740-1 (1-173) x ABX38978 (1-380)

Qy         54 ArgThrValLeuAsp---SerGlyIleSerGluValArgSerAspArgAspLysPheVal 72
              ||||||   ||||||   |||      ||||||:::|||   :::::::|||:::|||
Db         12 CGCACCCAGCTGGATTGCTCTGCCCCCTCAGAGATGCGCCTGGAGAAGGACAGATTCTCT 71

Qy         73 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 92
              :::   |||||:::   |||||||||||||||:::|||      ||||||      |||
Db         72 GTCAACCTGGACCTGACGCACTTCTCCCCAGAGGAACTCAAGGCCAAGGTGCTGGGAGAT 131

Qy         93 PheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArg 112
                 :::|||:::|||||||||||   ||||||||||||:::||||||:::|||||||||
Db        132 GTGATTGAGGTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGG 191

Qy        113 GluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 132
              ||||||||||:::||||||:::|||::::::||||||      |||::::::   |||
Db        192 GAGTTCCACAGGAAATACCGGATCCCAGCTGACGTGGACCCTCTCGCCATTACTTCATCC 251

Qy        133 LeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAla 152
              ||||||:::||||||   ||||||      ||||||:::   |||             |||
Db        252 CTGTCCTCTGATGGGGACCTCACTGTGAATGGACCAAGGAAACAG-----------GCC 299

Qy        153 ThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 167
                 :::      ||||||   ||||||::::::|||||||||||||||
```

```
Db        300 TCCGGCCCTGAGCGCACCATTCCCATAACCCGTGAAGAGAAGCCG 344

RESULT 48
AAC03894
ID   AAC03894 standard; cDNA; 695 BP.
XX
AC   AAC03894;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 3892.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   P-PSDB; AAG03888.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 3892; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors
XX
SQ   Sequence 695 BP; 183 A; 165 C; 162 G; 183 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              2.02e-32        Length:         695
Score:                  330.00          Matches:        62
Percent Similarity:     77.27%          Conservative:   23
Best Local Similarity:  56.36%          Mismatches:     19
Query Match:            36.03%          Indels:         6
DB:                     3               Gaps:           2

US-10-657-740-1 (1-173) x AAC03894 (1-695)

Qy         64 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 83
              :::|||   :::::::|||:::|||   :::   ||||||||||||||||||||||||||
Db        269 ATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTGGATGTGAAGCACTTCTCCCCAGAG 328

Qy         84 AspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArg 103
              :::|||   |||||||||       |||   :::|||:::||||||||||||   ||||||
Db        329 GAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAGGTGCATGGAAAACATGAAGAGCGC 388

Qy        104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
              ||||||:::||||||:::|||||||||||||||||||:::||||||:::|||::::::
Db        389 CAGGATGAACATGGTTTCATCTCCAGGGAGTTCCACAGGAAATACCGGATCCCAGCTGAT 448

Qy        124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
              ||||||          ::::::   |||||||||:::||||||:::||||||      |||
Db        449 GTAGACCCTCTCACCATTACTTCATCCCTGTCATCTGATGGGGTCCTCACTGTGAATGGA 508

Qy        144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
              |||:::   |||               :::      ||||||   ||||||:::::|||
Db        509 CCAAGGAAACAG-----------GTCTCTGGCCCTGAGCGCACCATTCCCATCACCCGT 556

Qy        164 GluGluLysPro------ThrSerAlaPro 171
              ||||||||||||      |||:::||||||
Db        557 GAAGAGAAGCCTGCTGTCACCGCAGCCCCC 586

RESULT 49
ACH16863
ID   ACH16863 standard; cDNA; 405 BP.
XX
AC   ACH16863;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult heart cDNA #1177.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 4075; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
```

```
XX
SQ   Sequence 405 BP; 50 A; 148 C; 137 G; 70 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.19e-30        Length:         405
Score:                  311.50          Matches:        60
Percent Similarity:     63.64%          Conservative:   24
Best Local Similarity:  45.45%          Mismatches:     41
Query Match:            34.01%          Indels:         7
DB:                     9               Gaps:           2

US-10-657-740-1 (1-173) x ACH16863 (1-405)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||      |||   :::|||          |||
Db         16 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 75

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db         76 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 135

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                  :::|||::::::|||||||||         |||           :::
Db        136 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 186

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                 :::::::::|||   :::|||       |||   :::   ||||||||||
Db        187 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 246

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        247 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 306

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::|||||:::::::::|||||||||||||||||||||
Db        307 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 366

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSer 130
              ||||||||      ||||||   :::|||::::::
Db        367 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACG 402

RESULT 50
ABL07843
ID   ABL07843 standard; cDNA; 719 BP.
XX
AC   ABL07843;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 18011.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB63740.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 18011; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 719 BP; 174 A; 208 C; 179 G; 158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.26e-29        Length:         719
Score:                  308.50          Matches:        69
Percent Similarity:     58.05%          Conservative:   32
Best Local Similarity:  39.66%          Mismatches:     54
Query Match:            33.68%          Indels:         19
DB:                     4               Gaps:           5

US-10-657-740-1 (1-173) x ABL07843 (1-719)

Qy          9 TrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePheGly 28
              |||:::          |||        |||||||||   ||||||   ||||||
Db         77 TGGTGGGATGAGCTCGACTTTCCAATGCGCACCTCGCGCCTCTTGGACCAGCACTTCGGT 136

Qy         29 GluGlyLeuPheGluTyrAspLeuLeuProPheLeu---------------------- 40
              :::||||||        |||||||:::      :::
Db        137 CAGGGACTGAAGAGGGATGACCTCATGTCCTCCGTGTGGAACTCCCGCCCCACAGTGCTG 196

Qy         41 ---SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSer 59
                    |||       :::   |||::::::      ||||||   :::         :::|||
Db        197 CGATCGGGATACCTGCGCCCCTGGCACACGAACAGCCTGCAGAAGCAG------GAGTCC 250

Qy         60 GlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHis 79
              .|||   |||   :::      |||  :::||||||   :::   ||||||||:::
Db        251 GGT---TCCACCCTCAACATCGACAGCGAGAAGTTCGAGGTCATTCTGGATGTGCAGCAG 307

Qy         80 PheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLys 99
              |||||||||   ::::::||||||||||||   |||   ||||||   :::   ||||||
Db        308 TTCTCGCCATCCGAGATCACCGTCAAGGTGGCCGACAAGTTCGTCATCGTGGAGGGAAAG 367

Qy        100 HisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArg 119
              |||   |||:::||||||:::||||||||||:::||||||:::|||   |||||||||:::
Db        368 CATGAGGAGAAGCAGGACGAGCACGGATATGTCTCCCGCCAGTTCTCCAGACGCTACCAG 427

Qy        120 LeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
              |||||||||:::|||:::        ::::::   |||||||||:::||||||:::|||
Db        428 CTGCCAAGTGATGTCAACCCCGACACGGTCACATCGTCCCTCTCCTCCGATGGCCTGCTG 487

Qy        140 ThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIle 159
              |||         |||              |||                |||||   :::
Db        488 ACCATCAAGGCGCCA------ATGAAGGCGCTGCCTCCTCCGCAGACTGAACGCCTTGTG 541

Qy        160 ProValSerArgGluGluLysProThrSerAlaProSerSer 173
                 :::::::::             :::   |||||||||
Db        542 CAGATCACACAG--------------ACTGGTCCCTCCTCC 568


Search completed: May 30, 2005, 05:14:43
Job time : 550.598 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2005, 04:24:56 ; Search time 170.078 Seconds
(without alignments)
1664.391 Million cell updates/sec

Title: US-10-657-740-1
Perfect score: 916
Sequence: 1 MDVTIQHPWFKRTLGPFYPS.....HAERAIPVSREEKPTSAPSS 173

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10657740/runat_27052005_165254_3343/app_query.fasta_1.590
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=150 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10657740_@CGN_1_1_219_@runat_27052005_165254_3343 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	916	100.0	1114	4	US-09-949-016-5220	Sequence 5220, Ap
2	489	53.4	691	4	US-09-949-016-2545	Sequence 2545, Ap
3	489	53.4	856	4	US-09-513-999C-3891	Sequence 3891, Ap
4	489	53.4	893	4	US-09-513-999C-14942	Sequence 14942, A
5	489	53.4	911	4	US-09-513-999C-14938	Sequence 14938, A
6	489	53.4	927	4	US-09-513-999C-14936	Sequence 14936, A
7	489	53.4	942	4	US-09-513-999C-3890	Sequence 3890, Ap
8	457.5	49.9	913	4	US-09-513-999C-14937	Sequence 14937, A
9	429	46.8	7775	4	US-09-949-016-16962	Sequence 16962, A
10	406.5	44.4	826	4	US-09-621-976-3172	Sequence 3172, Ap
11	330	36.0	695	4	US-09-513-999C-3892	Sequence 3892, Ap
12	289.5	31.6	789	4	US-09-023-655-1300	Sequence 1300, Ap
13	289.5	31.6	845	4	US-09-949-016-4992	Sequence 4992, Ap
14	289.5	31.6	1380	3	US-09-513-783A-169	Sequence 169, App
15	283	30.9	1379	3	US-09-553-498-5	Sequence 5, Appli
16	283	30.9	1379	3	US-09-618-869-5	Sequence 5, Appli
17	230	25.1	601	4	US-09-949-016-182120	Sequence 182120,
18	188	20.5	7125	4	US-09-949-016-14287	Sequence 14287, A
19	182.5	19.9	393	4	US-09-513-999C-10503	Sequence 10503, A
20	171.5	18.7	1627	2	US-08-900-407-2	Sequence 2, Appli
21	163	17.8	509	4	US-09-513-999C-1048	Sequence 1048, Ap
22	163	17.8	546	4	US-09-513-999C-1051	Sequence 1051, Ap
23	163	17.8	564	4	US-09-513-999C-1047	Sequence 1047, Ap
24	163	17.8	580	4	US-09-513-999C-1049	Sequence 1049, Ap
25	163	17.8	595	4	US-09-513-999C-1046	Sequence 1046, Ap
26	158	17.2	601	4	US-09-949-016-91317	Sequence 91317, A
27	158	17.2	763	4	US-09-949-016-2623	Sequence 2623, Ap
28	158	17.2	4763	4	US-09-949-016-14365	Sequence 14365, A
29	149.5	16.3	333	4	US-09-513-999C-10502	Sequence 10502, A
30	147	16.0	5688	4	US-09-949-016-16734	Sequence 16734, A
c 31	142	15.5	601	4	US-09-949-016-89647	Sequence 89647, A
32	140.5	15.3	599	4	US-09-270-767-563	Sequence 563, App
33	140.5	15.3	599	4	US-09-270-767-15845	Sequence 15845, A
34	131.5	14.4	566	4	US-09-513-999C-1052	Sequence 1052, Ap
35	125.5	13.7	495	4	US-09-669-751-69	Sequence 69, Appl
c 36	118	12.9	601	4	US-09-949-016-89648	Sequence 89648, A
c 37	115	12.6	12214	4	US-09-949-016-17284	Sequence 17284, A
38	110.5	12.1	719	4	US-09-620-312D-669	Sequence 669, App
c 39	110	12.0	9615	4	US-09-949-016-17282	Sequence 17282, A
c 40	107	11.7	7125	4	US-09-949-016-14287	Sequence 14287, A
41	97	10.6	601	4	US-09-949-016-175645	Sequence 175645,
42	96	10.5	420	4	US-09-107-532A-670	Sequence 670, App
43	93.5	10.2	652	1	US-07-748-761-2	Sequence 2, Appli
44	92	10.0	465	4	US-09-252-991A-12236	Sequence 12236, A
c 45	92	10.0	522	4	US-09-252-991A-12373	Sequence 12373, A
46	92	10.0	576	4	US-09-489-039A-3463	Sequence 3463, Ap
c 47	88	9.6	168971	4	US-09-949-016-13807	Sequence 13807, A
48	86	9.4	489	4	US-09-902-540-6436	Sequence 6436, Ap
49	86	9.4	3626	4	US-09-902-540-444	Sequence 444, App
c 50	84	9.2	49487	4	US-09-949-016-11770	Sequence 11770, A
51	84	9.2	57751	4	US-09-949-016-13631	Sequence 13631, A
52	82	9.0	20757	4	US-09-902-540-1189	Sequence 1189, Ap
53	79.5	8.7	1350	4	US-09-583-110-2262	Sequence 2262, Ap
54	79.5	8.7	1353	4	US-09-107-433-2322	Sequence 2322, Ap
55	79.5	8.7	9370	1	US-08-320-559-27	Sequence 27, Appl
56	79.5	8.7	9370	3	US-08-545-860D-27	Sequence 27, Appl
57	79.5	8.7	9370	5	PCT-US94-04496-27	Sequence 27, Appl
58	79.5	8.7	9391	1	US-08-320-559-25	Sequence 25, Appl
59	79.5	8.7	9391	3	US-08-545-860D-25	Sequence 25, Appl
60	79.5	8.7	9391	5	PCT-US94-04496-25	Sequence 25, Appl
61	77	8.4	1830121	4	US-09-557-884-1	Sequence 1, Appli
62	77	8.4	1830121	4	US-09-643-990A-1	Sequence 1, Appli
63	76.5	8.4	510	4	US-09-489-039A-3514	Sequence 3514, Ap
64	76.5	8.4	1008	4	US-09-969-532-19	Sequence 19, Appl
65	76.5	8.4	1041	4	US-09-969-532-17	Sequence 17, Appl
66	76.5	8.4	1203	4	US-09-602-787A-149	Sequence 149, App
67	76.5	8.4	1701	4	US-09-969-532-3	Sequence 3, Appli
68	76.5	8.4	1734	4	US-09-969-532-1	Sequence 1, Appli
69	76.5	8.4	2010	4	US-09-969-532-27	Sequence 27, Appl
70	76.5	8.4	2043	4	US-09-969-532-25	Sequence 25, Appl
c 71	76.5	8.4	2055	4	US-10-029-180-81	Sequence 81, Appl
72	76.5	8.4	2703	4	US-09-969-532-11	Sequence 11, Appl
73	76.5	8.4	2736	4	US-09-969-532-9	Sequence 9, Appli
74	76.5	8.4	3411	4	US-09-969-532-33	Sequence 33, Appl
75	76	8.3	9432	1	US-08-277-231A-1	Sequence 1, Appli
76	76	8.3	9432	2	US-08-473-750-4	Sequence 4, Appli
77	76	8.3	9432	2	US-08-477-326-4	Sequence 4, Appli
78	75.5	8.2	966	4	US-09-969-532-23	Sequence 23, Appl
79	75.5	8.2	999	4	US-09-969-532-21	Sequence 21, Appl
80	75.5	8.2	1659	4	US-09-969-532-7	Sequence 7, Appli
81	75.5	8.2	1692	4	US-09-969-532-5	Sequence 5, Appli
82	75.5	8.2	1968	4	US-09-969-532-31	Sequence 31, Appl
83	75.5	8.2	2001	4	US-09-969-532-29	Sequence 29, Appl
84	75.5	8.2	2661	4	US-09-969-532-15	Sequence 15, Appl
85	75.5	8.2	2694	4	US-09-969-532-13	Sequence 13, Appl

	86	75	8.2	360	4	US-09-543-681A-3563 Sequence 3563, Ap
c	87	74.5	8.1	34030	4	US-09-949-016-12248 Sequence 12248, A
c	88	74.5	8.1	34031	4	US-09-949-016-13540 Sequence 13540, A
	89	73.5	8.0	712	4	US-09-835-909A-1 Sequence 1, Appli
	90	73	8.0	633	4	US-09-252-991A-7328 Sequence 7328, Ap
	91	73	8.0	771	4	US-09-252-991A-7403 Sequence 7403, Ap
	92	73	8.0	1254	4	US-09-252-991A-7470 Sequence 7470, Ap
c	93	73	8.0	2103	4	US-09-252-991A-7175 Sequence 7175, Ap
	94	73	8.0	3402	1	US-08-480-917-1 Sequence 1, Appli
	95	73	8.0	3402	3	US-09-138-736-1 Sequence 1, Appli
	96	73	8.0	3402	3	US-08-988-242-1 Sequence 1, Appli
	97	73	8.0	62354	4	US-09-949-016-16188 Sequence 16188, A
c	98	72.5	7.9	978	4	US-09-248-796A-826 Sequence 826, App
	99	72.5	7.9	3000	2	US-08-896-344A-1 Sequence 1, Appli
	100	72.5	7.9	3000	3	US-09-360-682A-1 Sequence 1, Appli
	101	72.5	7.9	50341	1	US-08-247-901C-1 Sequence 1, Appli
	102	72.5	7.9	50341	2	US-09-075-904-1 Sequence 1, Appli
	103	72.5	7.9	52297	3	US-09-426-436-1 Sequence 1, Appli
	104	72.5	7.9	52297	3	US-08-705-557-1 Sequence 1, Appli
c	105	72.5	7.9	139936	4	US-09-949-016-11782 Sequence 11782, A
c	106	72.5	7.9	139952	4	US-09-949-016-13280 Sequence 13280, A
	107	72	7.9	2041	4	US-09-465-559-13 Sequence 13, Appl
c	108	72	7.9	62908	4	US-09-949-016-17554 Sequence 17554, A
	109	71.5	7.8	495	4	US-09-252-991A-12170 Sequence 12170, A
c	110	71.5	7.8	1354	3	US-08-456-200B-4 Sequence 4, Appli
	111	71.5	7.8	2881	4	US-09-949-016-5391 Sequence 5391, Ap
c	112	71.5	7.8	3563	3	US-09-041-886-20 Sequence 20, Appl
c	113	71.5	7.8	3596	2	US-08-779-801-5 Sequence 5, Appli
c	114	71.5	7.8	3596	3	US-09-298-441-5 Sequence 5, Appli
c	115	71.5	7.8	3632	2	US-08-779-801-3 Sequence 3, Appli
c	116	71.5	7.8	3632	2	US-08-779-801-4 Sequence 4, Appli
c	117	71.5	7.8	3632	3	US-09-298-441-3 Sequence 3, Appli
c	118	71.5	7.8	3632	3	US-09-298-441-4 Sequence 4, Appli
c	119	71.5	7.8	7791	2	US-08-149-097D-23 Sequence 23, Appl
c	120	71.5	7.8	7791	3	US-08-949-386-23 Sequence 23, Appl
c	121	71.5	7.8	7791	3	US-08-450-562-23 Sequence 23, Appl
c	122	71.5	7.8	7791	3	US-08-984-709A-23 Sequence 23, Appl
c	123	71.5	7.8	7791	3	US-08-450-272-23 Sequence 23, Appl
c	124	71.5	7.8	7791	4	US-08-450-273-23 Sequence 23, Appl
c	125	71.5	7.8	7808	2	US-08-149-097D-22 Sequence 22, Appl
c	126	71.5	7.8	7808	3	US-08-949-386-22 Sequence 22, Appl
c	127	71.5	7.8	7808	3	US-08-450-562-22 Sequence 22, Appl
c	128	71.5	7.8	7808	3	US-08-984-709A-22 Sequence 22, Appl
c	129	71.5	7.8	7808	3	US-08-450-272-22 Sequence 22, Appl
c	130	71.5	7.8	7808	4	US-08-450-273-22 Sequence 22, Appl
c	131	71.5	7.8	7815	4	US-09-949-016-3629 Sequence 3629, Ap
c	132	71.5	7.8	7815	4	US-09-949-016-3630 Sequence 3630, Ap
	133	71.5	7.8	70559	4	US-09-409-800B-1 Sequence 1, Appli
	134	71.5	7.8	1230230	4	US-09-438-185A-1 Sequence 1, Appli
	135	71	7.8	627	4	US-09-601-198-113 Sequence 113, App
c	136	71	7.8	1230	4	US-09-489-039A-6381 Sequence 6381, Ap
c	137	71	7.8	1278	4	US-09-902-540-8457 Sequence 8457, Ap
c	138	71	7.8	3762	4	US-09-435-376-1 Sequence 1, Appli
	139	71	7.8	8590	4	US-09-902-540-889 Sequence 889, App
	140	71	7.8	34794	4	US-09-713-678-39 Sequence 39, Appl
c	141	71	7.8	4403765	3	US-09-103-840A-2 Sequence 2, Appli
c	142	71	7.8	4411529	3	US-09-103-840A-1 Sequence 1, Appli
	143	70.5	7.7	585	4	US-09-248-796A-2425 Sequence 2425, Ap
	144	70.5	7.7	1605	3	US-09-124-541-3 Sequence 3, Appli
	145	70.5	7.7	1605	4	US-09-663-326-3 Sequence 3, Appli
c	146	70.5	7.7	5816	3	US-08-857-076-11 Sequence 11, Appl
	147	70.5	7.7	1230025	4	US-09-198-452A-1 Sequence 1, Appli
c	148	70	7.6	601	4	US-09-949-016-142871 Sequence 142871,
c	149	70	7.6	2247	4	US-09-252-991A-5125 Sequence 5125, Ap
	150	70	7.6	2952	4	US-09-252-991A-5096 Sequence 5096, Ap

ALIGNMENTS

RESULT 1
US-09-949-016-5220
; Sequence 5220, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5220
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5220

Alignment Scores:

Pred. No.:	1.7e-120	Length:	1114
Score:	916.00	Matches:	173
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x US-09-949-016-5220 (1-1114)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 129

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 189

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 249

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCTCTGAGGTTCGATCCGACCGGGACAAGTTCGTCATCTTCCTCGATGTGAAGCACTTC 309

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 TCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAGATCCACGGAAAGCAC 369

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTG 429

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 CCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACC 489

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCC 549

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||
Db        550 GTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 588
```

RESULT 2
US-09-949-016-2545
; Sequence 2545, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2545
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-2545

Alignment Scores:
Pred. No.:              7.89e-60        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     4               Gaps:           6

US-10-657-740-1 (1-173) x US-09-949-016-2545 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   ||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::|||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 3
US-09-513-999C-3891
; Sequence 3891, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 3891
;   LENGTH: 856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 4..753
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 51
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 117
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 118
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 159
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 160
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 182
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 183
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 16
;   OTHER INFORMATION: Xaa=Asp or Glu
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 38
;   OTHER INFORMATION: Xaa=His or Gln
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 39
;   OTHER INFORMATION: Xaa=Ala or Ser
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 53
;   OTHER INFORMATION: Xaa=Glu or Gln
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 60
;   OTHER INFORMATION: Xaa=Glu or Gly
US-09-513-999C-3891

Alignment Scores:
Pred. No.:              1.12e-59        Length:         856
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     4               Gaps:           6
```

```
US-10-657-740-1 (1-173) x US-09-513-999C-3891 (1-856)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        229 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 288

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        289 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 345

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        346 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 405

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        406 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 465

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        466 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 525

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        526 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 585

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::|||||         ::::::   |||||||||:::
Db        586 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 645

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||           :::
Db        646 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 693

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        694 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 747

RESULT 4
US-09-513-999C-14942
; Sequence 14942, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14942
;   LENGTH: 893
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-14942

Alignment Scores:
Pred. No.:              1.19e-59       Length:         893
Score:                  489.00         Matches:        97
Percent Similarity:     73.60%         Conservative:   34
Best Local Similarity:  54.49%         Mismatches:     35
Query Match:            53.38%         Indels:         12
DB:                     4              Gaps:           6

US-10-657-740-1 (1-173) x US-09-513-999C-14942 (1-893)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        266 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 325

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        326 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 382

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        383 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 442

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        443 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 502

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        503 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 562

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        563 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 622

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::|||||         ::::::   |||||||||:::
Db        623 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 682

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||           :::
Db        683 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 730

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        731 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 784

RESULT 5
US-09-513-999C-14938
; Sequence 14938, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14938
;   LENGTH: 911
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-14938

Alignment Scores:
Pred. No.:              1.23e-59       Length:         911
Score:                  489.00         Matches:        97
Percent Similarity:     73.60%         Conservative:   34
Best Local Similarity:  54.49%         Mismatches:     35
Query Match:            53.38%         Indels:         12
DB:                     4              Gaps:           6

US-10-657-740-1 (1-173) x US-09-513-999C-14938 (1-911)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        284 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 343

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
```

```
Db        344 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 400

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55

Db        401 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 460

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75

Db        461 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 520

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95

Db        521 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 580

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115

Db        581 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 640

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135

Db        641 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 700

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155

Db        701 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 748

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171

Db        749 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 802
```

RESULT 6
US-09-513-999C-14936
; Sequence 14936, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14936
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-14936

Alignment Scores:

Pred. No.:	1.27e-59	Length:	927
Score:	489.00	Matches:	97
Percent Similarity:	73.60%	Conservative:	34
Best Local Similarity:	54.49%	Mismatches:	35
Query Match:	53.38%	Indels:	12
DB:	4	Gaps:	6

US-10-657-740-1 (1-173) x US-09-513-999C-14936 (1-927)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19

Db        300 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 359

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39

Db        360 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 416

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55

Db        417 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 476

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75

Db        477 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 536

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95

Db        537 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 596

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115

Db        597 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 656

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135

Db        657 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 716

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155

Db        717 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 764

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171

Db        765 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 818
```

RESULT 7
US-09-513-999C-3890
; Sequence 3890, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3890
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 315..839
US-09-513-999C-3890

Alignment Scores:

Pred. No.:	1.3e-59	Length:	942
Score:	489.00	Matches:	97
Percent Similarity:	73.60%	Conservative:	34
Best Local Similarity:	54.49%	Mismatches:	35
Query Match:	53.38%	Indels:	12
DB:	4	Gaps:	6

US-10-657-740-1 (1-173) x US-09-513-999C-3890 (1-942)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19

Db        315 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 374

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39

Db        375 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 431

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
```

```
Db        432 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 491

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        492 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 551

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        552 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 611

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::||||||:::||||||||||||||||||||
Db        612 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 671

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        672 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 731

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              |||||||:::||||||      ||||||:::   |||               :::
Db        732 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 779

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
Db        780 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 833

RESULT 8
US-09-513-999C-14937
; Sequence 14937, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14937
;   LENGTH: 913
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 203
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 204
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 330
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 372
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 390
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 394
;   OTHER INFORMATION: s=g or c
US-09-513-999C-14937

Alignment Scores:
Pred. No.:              3.91e-55        Length:         913
Score:                  457.50          Matches:        94
Percent Similarity:     71.91%          Conservative:   34
Best Local Similarity:  52.81%          Mismatches:     33
Query Match:            49.95%          Indels:         17
DB:                     4               Gaps:           7

US-10-657-740-1 (1-173) x US-09-513-999C-14937 (1-913)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              |||||:::   |||   ||||||:::                  ||||||:::   |||
Db        301 ATGGACATCGCCATCCACCACCCCTTC--------------TTKCCTTTCCACTCCCCC 345

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   |||      |||
Db        346 AGCCGCCTCTTTGACCAGTTCTTCGGRGAGCACCTGTTGGAGTCKGATSTTTTCCCG--- 402

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::||||||::|||   :::      |||      |||      :::
Db        403 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 462

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        463 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 522

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        523 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 582

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::||||||:::||||||||||||||||||||
Db        583 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 642

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        643 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 702

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              |||||||:::||||||      ||||||:::   |||               :::
Db        703 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 750

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
Db        751 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 804

RESULT 9
US-09-949-016-16962
; Sequence 16962, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16962
;   LENGTH: 7775
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16962

Alignment Scores:
Pred. No.:              1.48e-49        Length:         7775
```

Score:	429.00	Matches:	172
Percent Similarity:	16.23%	Conservative:	0
Best Local Similarity:	16.23%	Mismatches:	1
Query Match:	46.83%	Indels:	888
DB:	4	Gaps:	2

US-10-657-740-1 (1-173) x US-09-949-016-16962 (1-7775)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2070 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 2129

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 2189

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 2249

Qy         61 IleSerGluVal---------------------------------------------- 64
              |||||||||
Db       2250 ATCTCTGAGGT-AAGACGTGGCCCGTGGTGCTGGCCTCTCCTCGCTGCTCAGAGGGTGGT 2308

Qy         64 ------------------------------------------------------------ 64

Db       2309 GGCCTCGGTGGGTGGAGAGCGATGGACTCTGGTCTTGCTCCGTCAGGCAGGTGGCCTCGT 2368

Qy         64 ------------------------------------------------------------ 64

Db       2369 CCCACTTCATCCCCTTGCAGAGGCTGGGCGAGAGCCTGTGTCCCCACTGCAGCCACGTGG 2428

Qy         64 ------------------------------------------------------------ 64

Db       2429 CAGAGCTTCCCCTGGCACTGGGGAGAGGGTGGACAAGGGAGCAGCCTGAATCCACCTTTG 2488

Qy         64 ------------------------------------------------------------ 64

Db       2489 CTTTCCTCCATCAGCTCATGACCCATGTGTGCTTTGTAAGGCACCAGCCACATACTGGAA 2548

Qy         64 ------------------------------------------------------------ 64

Db       2549 ACCCCAAGAGCAGCCCATCCAGGCATGCGTGGTGGCGAATGCCAGCTCCCGGGTTCCTCT 2608

Qy         64 ------------------------------------------------------------ 64

Db       2609 GGTCTCCTGAGTCCCGGAGACCTGGGAGCAGGTGGGGGTCATAGTCCTGAAAGCCAGAGA 2668

Qy         64 ------------------------------------------------------------ 64

Db       2669 GCAGGGCGTTCCTAGCACCTCCTCCAATGAGCTCGGCCTGCCCACGGCTAGCAAAGCTCT 2728

Qy         64 ------------------------------------------------------------ 64

Db       2729 TGGCAAGTTTACTTAGGTGCCCTGCCAAGGCTAAAAGGACAGGCAATGGACGCCCCCCCC 2788

Qy         64 ------------------------------------------------------------ 64

Db       2789 CCCCACCCAACCACAGGCCTCCTCTCTGAGCCACGGGTGAGCGGTGCAGGTTCTGCTGTT 2848

Qy         64 ------------------------------------------------------------ 64

Db       2849 CTGGAGGGCCTGAGTCCCACCCAGCACCTCATAAACAGGGTCCTCCCCAGGGCTGCTGCA 2908

Qy         64 ------------------------------------------------------------ 64

Db       2909 GTAGGCATCAATGCCAGGGTGCAAAATGCCTCAGGGAGCCAAGGCTGAGCCAGGGGAGTG 2968

Qy         64 ------------------------------------------------------------ 64

Db       2969 AGAAGGAGCATGTGGAAGTGCGTTTTGGAGAGGCAGCTGCGCAGGCTGTCAGCAGGCTCC 3028

Qy         64 ------------------------------------------------------------ 64

Db       3029 GGCCGCTTCTATAGACAGCATGACACCAAGGGCAGTGACCTCATTCCACAGGCTGAGTCC 3088

Qy         64 ------------------------------------------------------------ 64

Db       3089 AGCCAGCCAGCCAAGCATCACCAGCCAGACGATTGACCCTAACGGACCAACCAACCCGTA 3148

Qy         64 ------------------------------------------------------------ 64

Db       3149 ACGACCCCTCCTACCATAACCAGTAGCCAGCCAGCCCATAACCAGCCAACTTATCTATAA 3208

Qy         64 ------------------------------------------------------------ 64

Db       3209 CCAGCCACCTGACCATAGCCAAACAACCAGCCGGCCCACCAGTAGCATTCAGCCCCTCAG 3268

Qy         64 ------------------------------------------------------------ 64

Db       3269 CTGGCCCTGAGGGTTTGGAGACAGGTCGAGGGTCATGCCTGTCTGTCCAGGAGACAGTCA 3328

Qy         64 ------------------------------------------------------------ 64

Db       3329 CAGGCCCCCGAAAGCTCTGCCCCACTTGGTGTGTGGGAGAAGAGGCCGGCAGGTGACCGA 3388

Qy         64 ------------------------------------------------------------ 64

Db       3389 AGCATCTCTGTTCTGATAACCGGGACCCGCCCTGTCTCTGCCAACCCCAGCAGGGACGGC 3448

Qy         65 ----------------------------------------------ArgSerAspArgAspLys 70
                                                            ||||||||||||||||||
Db       3449 ACCCTCTGGGCAGCTCCACATGGCACGTTTGGATTTCAGGTTCGATCCGACCGGGACAAG 3508

Qy         71 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 TTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGACCTCACCGTGAAGGTGCAG 3568

Qy         91 AspAspPheValGluIleHisGlyLysHisAsnGluArg---------------------- 103
              ||||||||||||||||||||||||||||||||||||||||
Db       3569 GACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAGGTGAGCCCAGGCACTGAG 3628

Qy        103 ------------------------------------------------------------ 103

Db       3629 AGGTGGGAGAGGGGGGCGAGTTGGGCGCGAGGACAAGGGGGTCACGGCGGGCACGACCGG 3688

Qy        103 ------------------------------------------------------------ 103

Db       3689 GCCTGCACACCTGCACCATGCCTTCAACCCTGGGAGAGGGACGCTCTCCAGGGGACCCCG 3748

Qy        103 ------------------------------------------------------------ 103

Db       3749 AATCAGGCCTGGCTTTTCCCCAAGGGAGGGGCCATGCCCACCTGAGCACAGCCAGCCCCT 3808

Qy        103 ------------------------------------------------------------ 103

Db       3809 CCCGGTGACAGAGGTCACCATTCCCGAGCTAATGTGGCTCAGGGATCCAGGTTAGGGTCC 3868

Qy        103 ------------------------------------------------------------ 103

Db       3869 CTTCCCGGGCTGCACCCAGCCGTCGCCAGCTCCATCCCTGTCACCTGGATGCCAGGGTGG 3928

Qy        103 ------------------------------------------------------------ 103

Db       3929 TCTTAGAAAGAACCCCAGGAAGTGGGAGTGCCCCGGGTGGCCGCCTCCTAGCCAGTGTAC 3988

Qy        103 ------------------------------------------------------------ 103

Db       3989 ATCTTCACATGAACCCTACCTGAGGAAGCCAGTCCCCGACGGCATAGCTGCATCCGCTTG 4048

Qy        103 ------------------------------------------------------------ 103

Db       4049 GAATGCTTTACAGGCATTGACACCTTCGCCTCACAGCAGCACTTTGGAACCAGTGTCCTC 4108

Qy        103 ------------------------------------------------------------ 103
```

```
Db       4109 ATTATTCCAGGGCACGGCTGGGGAACAAGGGGGTCCTCAGCCTGCTGGGTCCCACAGCTA 4168
Qy        103 ------------------------------------------------------------ 103
Db       4169 GTACCGGGCAGGTGGACGGGAGCTTCTCCCCACAGTCACCCTGATGCCCCGCTCTTGCTC 4228
Qy        103 ------------------------------------------------------------ 103
Db       4229 GGCTGGAGGCCTCGGATCTCCGTGGTGTTGAGGGAGCCGGGGCACTGGAGCCCTGGTGAC 4288
Qy        103 ------------------------------------------------------------ 103
Db       4289 CTGCATCTCCTGGCGGAGCCGGGAAGAGCTCATGGACTGTCACAGATGGACAGTGCCCCG 4348
Qy        103 ------------------------------------------------------------ 103
Db       4349 CGGGGGCTGGAGAGCAGAGTGGGGCTGGAAGGTGGAACTCTTAGCCAAAGTCTTGGTTTC 4408
Qy        103 ------------------------------------------------------------ 103
Db       4409 TTTTGGCCAGGGTCCTCTTTCAATGGCTGGAGAAGGTGGTGCTGGGGGGTGAACGCTGAC 4468
Qy        103 ------------------------------------------------------------ 103
Db       4469 CTCCTCATGTGCTGCCCCTCCCTCGCCTGGGCCCGGTAAAGCCCCCACGTAGCCCCAGCC 4528
Qy        103 ------------------------------------------------------------ 103
Db       4529 AGCCTGGAACATGCTTCCTGAGCTCCCAGCTCTTGGTCTTTGCACCCAGTGGAGGAGGAG 4588
Qy        103 ------------------------------------------------------------ 103
Db       4589 GTCAGCCCAGGGAGCTGAGTCTGCGGTTTAGGGCGTCCAGGGGACGTGGAAGCATGTGGG 4648
Qy        103 ------------------------------------------------------------ 103
Db       4649 TCGTCTGGCCACATTAGGTAGGGCTGCAGAGACCTGGGCTAGAGCAGTCCTGCGGGGTCT 4708
Qy        103 ------------------------------------------------------------ 103
Db       4709 GGAAGGGGAAGACTGGCTGAGGTGCGGGGCCTGGTCTGGAATGATCCTGCGATTTTGGAG 4768
Qy        103 ------------------------------------------------------------ 103
Db       4769 TGAAGCCATGGAGCGGGAAGAGACAACCCCCCGCGGGGAATAGCCCGGCAAGTGGCCACG 4828
Qy        103 ------------------------------------------------------------ 103
Db       4829 AGGCCAGGCTGAGGTCCAGAGAAGCAGGGGCATGAATCCATAAATCCCAGGGGGCCTGGC 4888
Qy        103 ------------------------------------------------------------ 103
Db       4889 CATGGGATGTGCTGGCTGCACCCGGCCCCTGTGAGAGCCCCCGCAGGCTGGCCCCCTTCT 4948
Qy        103 ------------------------------------------------------------ 103
Db       4949 GCAGTCAGTGGGGCTGGGGCAGCTTCTCTGGCATGGGGCGAGGCAGCCGCCTGCACAGTG 5008
Qy        104 ----------------------------GlnAspAspHisGlyTyrIleSerArgGlu 113
                                           ||||||||||||||||||||||||||||||
Db       5009 GCCCCCCTGACTGTGCGCCCCCACCCTCTCCAGGACGACCACGGCTACATTTCCCGTGAG 5068
Qy        114 PheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeu 133
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5069 TTCCACCGCCGCTACCGCCTGCCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTG 5128
Qy        134 SerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThr 153
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5129 TCTGCCGATGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACC 5188
Qy        154 HisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5189 CACGCCGAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 5248
```

```
RESULT 10
US-09-621-976-3172
; Sequence 3172, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 3172
;   LENGTH: 826
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 557..736
US-09-621-976-3172

Alignment Scores:
Pred. No.:              6.43e-48        Length:         826
Score:                  406.50          Matches:        93
Percent Similarity:     68.72%          Conservative:   30
Best Local Similarity:  51.96%          Mismatches:     37
Query Match:            44.38%          Indels:         19
DB:                     4               Gaps:           5

US-10-657-740-1 (1-173) x US-09-621-976-3172 (1-826)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeu-GlyProPheTyr---Pr 19
              ||||||:::   |||   ||||||      ::::::   |||     ||||||:::   ||
Db        267 ATGGACATCGCCATCCACCACCCCTTGATCCGCCAGCSCCTTCTTTCCTTTCCACTCCCC 326

Qy         19 oSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAsp----------- 35
              ||||||||||||||||||||||||||||||   |||   |||   |||
Db        327 CAGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATACTATTTCCCG 386

Qy         36 ----LeuLeuProPheLeuSerSerThrIleSer------ProTyrTyrArgGlnSerLe 52
                  |||||||||   :::   ||||||           ||||||   |||
Db        387 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTACCTGCGGGCACCCAG 446

Qy         52 uPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheVa 72
              :::           |||:::|||:::||||||:::|||   :::::::|||:::|||
Db        447 CTGG---------TTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTC 497

Qy         72 lIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAs 92
              :::   ||||||||||||||||||||||||||:::|||   |||||||||      ||
Db        498 TGTCAACCTGGATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGA 557

Qy         92 pPhe-ValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerA 112
              |    :::|||:::|||||||||||   ||||||||||:::||||||:::|||||||
Db        558 TGTCGATTGAGGTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCA 617

Qy        112 rgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysS 132
              |||||||||||||:::||||||:::|||:::::|||||      ::::::   |
Db        618 GGGAGTTCCACAGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCAT 677

Qy        132 erLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspA 152
              ||||||||:::||||||:::||||||       ||||||:::   |||
Db        678 CCCTGTCATCTGATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------G 725

Qy        152 laThr-HisAlaGluArgAlaIleProValSer-ArgGluGluLysPro 167
              ,     :::       ||||||   ||||||:::::: ||||||||||||||
Db        726 TCTCTGGGCCCTGAGCGCACCATTCCCATCACCCCGTGAAGAGAAGCCT 774

RESULT 11
US-09-513-999C-3892
; Sequence 3892, Application US/09513999C
```

```
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 3892
;   LENGTH: 695
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 269..592
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 44
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45
;   OTHER INFORMATION: s=g or c
US-09-513-999C-3892

Alignment Scores:
Pred. No.:              4.13e-37      Length:         695
Score:                  330.00        Matches:        62
Percent Similarity:     77.27%        Conservative:   23
Best Local Similarity:  56.36%        Mismatches:     19
Query Match:            36.03%        Indels:         6
DB:                     4             Gaps:           2

US-10-657-740-1 (1-173) x US-09-513-999C-3892 (1-695)

Qy          64 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 83
               :::|||   :::::::|||:::|||   :::   ||||||||||||||||||||||||
Db         269 ATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTGGATGTGAAGCACTTCTCCCCAGAG 328

Qy          84 AspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArg 103
               :::|||   |||||||||       |||   :::|||:::|||||||||||   ||||||
Db         329 GAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAGGTGCATGGAAAACATGAAGAGCGC 388

Qy         104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
               ||||||:::|||||||:::|||||||||||||||||||||:::|||||||:::|||:::::
Db         389 CAGGATGAACATGGTTTCATCTCCAGGGAGTTCCACAGGAAATACCGGATCCCAGCTGAT 448

Qy         124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
               ||||||        ::::::   |||||||||:::||||||:::||||||       |||
Db         449 GTAGACCCTCTCACCATTACTTCATCCCTGTCATCTGATGGGGTCCTCACTGTGAATGGA 508

Qy         144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
               |||:::   |||              :::      ||||||   ||||||::::::|||
Db         509 CCAAGGAAACAG-----------GTCTCTGGCCCTGAGCGCACCATTCCCATCACCCGT 556

Qy         164 GluGluLysPro------ThrSerAlaPro 171
               ||||||||||||      |||:::||||||
Db         557 GAAGAGAAGCCTGCTGTCACCGCAGCCCCC 586

RESULT 12
US-09-023-655-1300
; Sequence 1300, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1300:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  789 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g32477
US-09-023-655-1300

Alignment Scores:
Pred. No.:              3.1e-31       Length:         789
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     4             Gaps:           6

US-10-657-740-1 (1-173) x US-09-023-655-1300 (1-789)

Qy          10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
               |||                  ||||||                  :::|||||   ||||||
Db          74 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 133

Qy          22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
               ||||||||||||   ||||||                              ||||||   |||
Db         134 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 169

Qy          40 ------------------------------------------------------------ 40

Db         170 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 229

Qy          41 -----------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                                     :::      :::|||   |||   ::::::|||   |||
Db         230 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 289

Qy          55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
               |||   ||||||:::||||||:::|||          |||:::::   :::
Db         290 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 346

Qy          75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
               |||||||||   ||||||:::|||::::::||||||||||||   :::|||      |||
```

```
Db        347 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 406

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114

Db        407 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 466

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134

Db        467 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 526

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154

Db        527 CCTGAGGGCACACTGACCGTGGA-GGCCCC-------------------------CAT 558

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168

Db        559 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 600

RESULT 13
US-09-949-016-4992
; Sequence 4992, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4992
;   LENGTH: 845
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-4992

Alignment Scores:
Pred. No.:              3.46e-31      Length:         845
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     4             Gaps:           6

US-10-657-740-1 (1-173) x US-09-949-016-4992 (1-845)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21

Db        129 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 188

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40

Db        189 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 224

Qy         40 ------------------------------------------------------------ 40

Db        225 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 284

Qy         41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54

Db        285 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 344

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74

Db        345 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 401

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94

Db        402 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 461

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114

Db        462 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 521

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134

Db        522 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 581

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154

Db        582 CCTGAGGGCACACTGACCGTGGA-GGCCCC-------------------------CAT 613

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168

Db        614 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 655

RESULT 14
US-09-513-783A-169
; Sequence 169, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
;  APPLICANT:  Giuliano, Kenneth A.
;  APPLICANT:  Kapur, Ravi
;  TITLE OF INVENTION:  A System for Cell Based Screening
;  FILE REFERENCE:  97-022-L1
;  CURRENT APPLICATION NUMBER: US/09/513,783A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS:  180
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 169
;   LENGTH: 1380
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GFP-HSP27
;   NAME/KEY: CDS
;   LOCATION: (1)..(1380)
US-09-513-783A-169

Alignment Scores:
Pred. No.:              7.65e-31      Length:         1380
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     3             Gaps:           6

US-10-657-740-1 (1-173) x US-09-513-783A-169 (1-1380)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21

Db        784 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 843

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40

Db        844 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 879

Qy         40 ------------------------------------------------------------ 40

Db        880 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 939

Qy         41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54

Db        940 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 999

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74

Db       1000 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 1056
```

```
Qy          75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
               |||||||||   ||||||:::|||::::::|||||||||||||   :::|||      |||
Db        1057 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 1116

Qy          95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
               ||||||   |||||||||   ||||||||||||:::|||||||||||||||||||   |||
Db        1117 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 1176

Qy         115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
               |||:::|||   ||||||      ||||||   :::   :::|||   |||||||||
Db        1177 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 1236

Qy         135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
               :::|||   ||||||      ||||||                          |||
Db        1237 CCTGAGGGCACACTGACCGTGGA-GGCCCC-------------------------CAT 1268

Qy         155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
               |||:::   :::      |||   |||:::      |||:::
Db        1269 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 1310
```

RESULT 15
US-09-553-498-5
; Sequence 5, Application US/09553498
; Patent No. 6309861
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: Process for the production of naturally folded and secreted prote
; FILE REFERENCE: Case 20379
; CURRENT APPLICATION NUMBER: US/09/553,498
; CURRENT FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: EP99107412.1
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 5
; LENGTH: 1379
; TYPE: DNA
; ORGANISM: E. coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (392)...(1090)
US-09-553-498-5

Alignment Scores:

Pred. No.:	6.48e-30	Length:	1379
Score:	283.00	Matches:	75
Percent Similarity:	51.34%	Conservative:	21
Best Local Similarity:	40.11%	Mismatches:	43
Query Match:	30.90%	Indels:	48
DB:	3	Gaps:	7

US-10-657-740-1 (1-173) x US-09-553-498-5 (1-1379)

```
Qy          16 ProPhe---------TyrPro------SerArgLeuPheAspGlnPhePheGly------ 28
               ||||||         ||||||      |||||||||||||||||   ||||||
Db         512 CCATTCCGGGACTGGTACCCTGCACACAGCCGCCTCTTCGATCAAGCTTTCGGGGTGCCC 571

Qy          28 ------------------------------------------------------------ 28

Db         572 CGGTTGCCCGATGAGTGGTCGCAGTGGTTCAGCGCCGCTGGGTGGCCCGGATACGTGCGC 631

Qy          29 --------------------GluGlyLeuPheGluTyrAspLeuLeuProPheLeuSer 41
                                   ||||||                        |||
Db         632 CCGCTGCCCGCCGCGACCGCCGAGGGC--------------------CCCGCGGCGGTG 670

Qy          42 SerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIle 61
               :::      :::|||   :::   :::::::|||   |||       |||   ||||||:::
Db         671 ACCCTGGCCGCACCAGCCTTCAGCCGAGCGCTCAACCGACAG---CTCAGCAGCGGGGTC 727

Qy          62 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 81
               ||||||:::|||         |||::::::   :::   |||||||||   ||||||:::
```

```
Db         728 TCGGAGATCCGACAGACGGCTGATCGCTGGCGCGTGTCCCTGGACGTCAACCACTTCGCT 787

Qy          82 ProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsn 101
               ||||||:::|||||||||||||   ::::::      |||||||||   |||||||||
Db         788 CCGGAGGAGCTCACAGTGAAGACCAAGGAAGGCGTGGTGGAGATCACTGGCAAGCACGAA 847

Qy         102 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
               ||||||||||||:::|||||||||||||||||   |||   |||:::|||   ||||||
Db         848 GAAAGGCAGGACGAACATGGCTACATCTCTCGGTGCTTCACCCGGAAATACACGCTCCCT 907

Qy         122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
                     ||||||   :::   :::|||   |||||||||   :::|||   ||||||
Db         908 CCAGGTGTGGACCCCACCCTAGTGTCCTCTTCCCTATCCCCTGAGGGCACACTTACCGTG 967

Qy         142 CysGly------ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIle 159
                           ||||||      |||                          ||||||      |||
Db         968 GAGGCTCCGTTGCCCAAAGCAGTCACG-----------CAGTCAGCGGAGATCACCATT 1015

Qy         160 ProValSerArgGluGluLys 166
               ||||||:::   |||   :::
Db        1016 CCGGTTACTTTCGAGGCCCGC 1036
```

RESULT 16
US-09-618-869-5
; Sequence 5, Application US/09618869
; Patent No. 6455279
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorthee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR
; TITLE OF INVENTION: CHAPERONES
; FILE REFERENCE: 20381
; CURRENT APPLICATION NUMBER: US/09/618,869
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: EP99114811.5
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1379
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (392)..(1090)
US-09-618-869-5

Alignment Scores:

Pred. No.:	6.48e-30	Length:	1379
Score:	283.00	Matches:	75
Percent Similarity:	51.34%	Conservative:	21
Best Local Similarity:	40.11%	Mismatches:	43
Query Match:	30.90%	Indels:	48
DB:	3	Gaps:	7

US-10-657-740-1 (1-173) x US-09-618-869-5 (1-1379)

```
Qy          16 ProPhe---------TyrPro------SerArgLeuPheAspGlnPhePheGly------ 28
               ||||||         ||||||      |||||||||||||||||   ||||||
Db         512 CCATTCCGGGACTGGTACCCTGCACACAGCCGCCTCTTCGATCAAGCTTTCGGGGTGCCC 571

Qy          28 ------------------------------------------------------------ 28

Db         572 CGGTTGCCCGATGAGTGGTCGCAGTGGTTCAGCGCCGCTGGGTGGCCCGGATACGTGCGC 631

Qy          29 --------------------GluGlyLeuPheGluTyrAspLeuLeuProPheLeuSer 41
                                   ||||||                        |||
Db         632 CCGCTGCCCGCCGCGACCGCCGAGGGC--------------------CCCGCGGCGGTG 670

Qy          42 SerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIle 61
```

```
Db        671 ACCCTGGCCGCACCAGCCTTCAGCCGAGCGCTCAACCGACAG---CTCAGCAGCGGGGTC 727

Qy         62 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 81
Db        728 TCGGAGATCCGACAGACGGCTGATCGCTGGCGCGTGTCCCTGGACGTCAACCACTTCGCT 787

Qy         82 ProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsn 101
Db        788 CCGGAGGAGCTCACAGTGAAGACCAAGGAAGGCGTGGTGGAGATCACTGGCAAGCACGAA 847

Qy        102 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
Db        848 GAAAGGCAGGACGAACATGGCTACATCTCTCGGTGCTTCACCCGGAAATACACGCTCCCT 907

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
Db        908 CCAGGTGTGGACCCCACCCTAGTGTCCTCTTCCCTATCCCCTGAGGGCACACTTACCGTG 967

Qy        142 CysGly------ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIle 159
Db        968 GAGGCTCCGTTGCCCAAAGCAGTCACG------------CAGTCAGCGGAGATCACCATT 1015

Qy        160 ProValSerArgGluGluLys 166
Db       1016 CCGGTTACTTTCGAGGCCCGC 1036
```

RESULT 17
US-09-949-016-182120
; Sequence 182120, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 182120
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-182120

Alignment Scores:

Pred. No.:	6.31e-23	Length:	601
Score:	230.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	25.11%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1 (1-173) x US-09-949-016-182120 (1-601)

```
Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
Db          2 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 61

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
Db         62 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 121

Qy         61 IleSerGluValArg 65
Db        122 ATCTCTGAGGTAAGA 136
```

RESULT 18
US-09-949-016-14287
; Sequence 14287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14287
; LENGTH: 7125
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(7125)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14287

Alignment Scores:

Pred. No.:	3.45e-15	Length:	7125
Score:	188.00	Matches:	37
Percent Similarity:	75.36%	Conservative:	15
Best Local Similarity:	53.62%	Mismatches:	11
Query Match:	20.52%	Indels:	6
DB:	4	Gaps:	2

US-10-657-740-1 (1-173) x US-09-949-016-14287 (1-7125)

```
Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
Db       4783 GATGAACATGGTTTCATCTCCAGGGAGTTCCACAGGAAATACCGGATCCCAGCTGATGTA 4842

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 144
Db       4843 GACCCTCTCACCATTACTTCATCCCTGTCATCTGATGGGGTCCTCACTGTGAATGGACCA 4902

Qy        145 LysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGlu 164
Db       4903 AGGAAACAG------------GTCTCTGGCCCTGAGCGCACCATTCCCATCACCCGTGAA 4950

Qy        165 GluLysPro------ThrSerAlaPro 171
Db       4951 GAGAAGCCTGCTGTCACCGCAGCCCCC 4977
```

RESULT 19
US-09-513-999C-10503
; Sequence 10503, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm

```
; SEQ ID NO 10503
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 75
;   OTHER INFORMATION: s=g or c
US-09-513-999C-10503

Alignment Scores:
Pred. No.:              1.94e-16      Length:         393
Score:                  182.50        Matches:        42
Percent Similarity:     60.23%        Conservative:   11
Best Local Similarity:  47.73%        Mismatches:     26
Query Match:            19.92%        Indels:         10
DB:                     4             Gaps:           1

US-10-657-740-1 (1-173) x US-09-513-999C-10503 (1-393)

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              |||      :::|||||||||||||   :::|||          ||||||   |||||||||
Db         36 AGCATGAACGAGCTGACGGTCAAGACCAAGGATGGCGTGSTGGAGATCACCGGCAAGCAC 95

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              |||||||||||:::|||||||||||||||||   |||   |||:::|||   |||
Db         96 GAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTCACGCGGAAATACACGCTG 155

Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              |||      ||||||   :::   :::|||   |||||||||   :::|||   ||||||
Db        156 CCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCCCCTGAGGGCACACTGACC 215

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              ||||||                                  ||||||:::   :::
Db        216 GTGGA-GGCCCC-------------------------CATGCCCAAGCTAGCCACGCA 247

Qy        161 ValSerArgGluGluLysProThr 168
              |||   |||:::      |||:::
Db        248 GTCCAACGAGATCACCATCCCAGT 271

RESULT 20
US-08-900-407-2
; Sequence 2, Application US/08900407
; Patent No. 5962262
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  NEW HUMAN HEAT SHOCK 27 LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/900,407
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0351 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1627 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  1
;    IMMEDIATE SOURCE:
;      LIBRARY:  ????
;      CLONE:  1362715
US-08-900-407-2

Alignment Scores:
Pred. No.:              7.19e-14      Length:         1627
Score:                  171.50        Matches:        45
Percent Similarity:     47.92%        Conservative:   24
Best Local Similarity:  31.25%        Mismatches:     58
Query Match:            18.72%        Indels:         17
DB:                     2             Gaps:           2

US-10-657-740-1 (1-173) x US-08-900-407-2 (1-1627)

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeu--------- 36
              |||||||||   |||      ||||||      |||      ||||||
Db        598 TCTCGCCTGCTGGATGATGGCTTTGGCATGGACCCCTTCCCAGACGACTTGACAGCCTCT 657

Qy         37 --------------LeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSer 51
                           ||||||   |||||||||                  |||
Db        658 TGGCCCGACTGGGCTCTGCCTCGTCTCTCCTCCGCCTGGCCAGGCACCCTAAGGTCGGGC 717

Qy         52 LeuPheArgThr--------------------------ValLeuAspSerGlyIleSer 62
              :::                                  |||      |||   :::
Db        718 ATGGTGCCCCGGGGCCCCACTGCCACCGCCAGGTTTGGGGTGCCTGCCGAGGGCAGGACC 777

Qy         63 GluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerPro 82
                        :::   :::   :::   ::::::|||      |||   |||
Db        778 CCCCCACCCTTCCCTGGGGAGCCCTGGAAAGTGTGTGTGAATGTGCACAGCTTCAAGCCA 837

Qy         83 GluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGlu 102
              |||:::|||   ||||||   :::|||   :::|||||||:::   |||||||||   |||
Db        838 GAGGAGTTGATGGTGAAGACCAAAGATGGATACGTGGAGGTGTCTGGCAAACATGAAGAG 897

Qy        103 ArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSer 122
              :::|||   :::   |||   :::|||:::   |||   ::::::   :::||||||:::
Db        898 AAACAGCAAGAAGGTGGCATTGTTTCTAAGAACTTCACAAAGAAAATCCAGCTTCCTGCA 957

Qy        123 AsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCys 142
              ||||||         :::      |||||||||   :::|||:::|||
Db        958 GAGGTGGATCCTGTGACAGTATTTGCCTCACTTTCCCCAGAGGGTCTGCTGATCATCGAA 1017

Qy        143 GlyProLysIle 146
              |||::::::
Db       1018 GCTCCCCAGGTC 1029

RESULT 21
US-09-513-999C-1048
; Sequence 1048, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
```

```
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1048
;   LENGTH: 509
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 4..507
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 51
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 117
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 118
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 159
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 160
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 182
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 183
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 16
;   OTHER INFORMATION: Xaa=Asp or Glu
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 38
;   OTHER INFORMATION: Xaa=His or Gln
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 39
;   OTHER INFORMATION: Xaa=Ala or Ser
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 53
;   OTHER INFORMATION: Xaa=Glu or Gln
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 60
;   OTHER INFORMATION: Xaa=Glu or Gly
US-09-513-999C-1048

Alignment Scores:
Pred. No.:              1.8e-13       Length:          509
Score:                  163.00        Matches:         36
Percent Similarity:     68.12%        Conservative:    11
Best Local Similarity:  52.17%        Mismatches:      16
Query Match:            17.79%        Indels:          6
DB:                     4             Gaps:            4

US-10-657-740-1 (1-173) x US-09-513-999C-1048 (1-509)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||       ||||||:::   |||
Db        229 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 288

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        289 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 345

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                |||::::::::::||||||:::|||   :::       |||       |||       :::
Db        346 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 405

Qy         56 ValLeuAspSerGlyIleSerGluVal 64
                  |||:::|||:::|||||||||
Db        406 TGGTTTGACACTGGACTCTCAGAGGTG 432

RESULT 22
US-09-513-999C-1051
; Sequence 1051, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1051
;   LENGTH: 546
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 266..544
US-09-513-999C-1051

Alignment Scores:
Pred. No.:              2.02e-13      Length:          546
Score:                  163.00        Matches:         36
Percent Similarity:     68.12%        Conservative:    11
Best Local Similarity:  52.17%        Mismatches:      16
Query Match:            17.79%        Indels:          6
DB:                     4             Gaps:            4

US-10-657-740-1 (1-173) x US-09-513-999C-1051 (1-546)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||       ||||||:::   |||
Db        266 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 325

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        326 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 382

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                |||::::::::::||||||:::|||   :::       |||       |||       :::
Db        383 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 442

Qy         56 ValLeuAspSerGlyIleSerGluVal 64
                  |||:::|||:::|||||||||
Db        443 TGGTTTGACACTGGACTCTCAGAGGTG 469

RESULT 23
US-09-513-999C-1047
; Sequence 1047, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
```

```
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1047
;   LENGTH: 564
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 284..562
US-09-513-999C-1047

Alignment Scores:
Pred. No.:              2.12e-13       Length:         564
Score:                  163.00         Matches:        36
Percent Similarity:     68.12%         Conservative:   11
Best Local Similarity:  52.17%         Mismatches:     16
Query Match:            17.79%         Indels:         6
DB:                     4              Gaps:           4

US-10-657-740-1 (1-173) x US-09-513-999C-1047 (1-564)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        284 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 343

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        344 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 400

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        401 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 460

Qy         56 ValLeuAspSerGlyIleSerGluVal 64
                 |||:::|||:::|||||||||
Db        461 TGGTTTGACACTGGACTCTCAGAGGTG 487

RESULT 24
US-09-513-999C-1049
; Sequence 1049, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1049
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 300..578
US-09-513-999C-1049

Alignment Scores:
Pred. No.:              2.22e-13       Length:         580
Score:                  163.00         Matches:        36
Percent Similarity:     68.12%         Conservative:   11
Best Local Similarity:  52.17%         Mismatches:     16
Query Match:            17.79%         Indels:         6
DB:                     4              Gaps:           4

US-10-657-740-1 (1-173) x US-09-513-999C-1049 (1-580)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        300 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 359

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        360 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 416

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        417 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 476

Qy         56 ValLeuAspSerGlyIleSerGluVal 64
                 |||:::|||:::|||||||||
Db        477 TGGTTTGACACTGGACTCTCAGAGGTG 503

RESULT 25
US-09-513-999C-1046
; Sequence 1046, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1046
;   LENGTH: 595
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 315..593
US-09-513-999C-1046

Alignment Scores:
Pred. No.:              2.32e-13       Length:         595
Score:                  163.00         Matches:        36
Percent Similarity:     68.12%         Conservative:   11
Best Local Similarity:  52.17%         Mismatches:     16
Query Match:            17.79%         Indels:         6
DB:                     4              Gaps:           4

US-10-657-740-1 (1-173) x US-09-513-999C-1046 (1-595)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db        315 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 374

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db        375 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 431

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        432 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 491

Qy         56 ValLeuAspSerGlyIleSerGluVal 64
                 |||:::|||:::|||||||||
```

```
Db        492 TGGTTTGACACTGGACTCTCAGAGGTG 518

RESULT 26
US-09-949-016-91317
; Sequence 91317, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91317
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-91317

Alignment Scores:
Pred. No.:              1.22e-12      Length:         601
Score:                  158.00        Matches:        31
Percent Similarity:     60.00%        Conservative:   14
Best Local Similarity:  41.33%        Mismatches:     30
Query Match:            17.25%        Indels:         0
DB:                     4             Gaps:           0

US-10-657-740-1 (1-173) x US-09-949-016-91317 (1-601)

Qy         65 ArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
              |||       :::       |||   |||   |||||||||       |||   |||||||||
Db        212 CGAGAAGGCAAATCCCACTTTCAGATCCTGCTGGACGTGGTCCAGTTCCTCCCTGAAGAC 271

Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              :::   ::::::      :::   ::::::   |||      :::|||      |||
Db        272 ATCATCATTCAGACCTTCGAAGGCTGGCTRCTGATAAAAGCACAACACGGAACCAGAATG 331

Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
              |||:::||||||:::|||||||||   |||   |||:::|||:::||||||      |||
Db        332 GATGAGCACGGTTTTATCTCAAGAAGCTTCACCCGACAGTACAAACTACCAGATGGCGTG 391

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
              :::       ||||||      |||      ||||||:::|||
Db        392 GAAATCAAAGATTTGTCTGCAGTCCTCTGTCATGATGGAATTTTG 436

RESULT 27
US-09-949-016-2623
; Sequence 2623, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2623
;   LENGTH: 763
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-2623

Alignment Scores:
Pred. No.:              1.79e-12      Length:         763
Score:                  158.00        Matches:        31
Percent Similarity:     60.00%        Conservative:   14
Best Local Similarity:  41.33%        Mismatches:     30
Query Match:            17.25%        Indels:         0
DB:                     4             Gaps:           0

US-10-657-740-1 (1-173) x US-09-949-016-2623 (1-763)

Qy         65 ArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
              |||       :::       |||   |||   |||||||||       |||   |||||||||
Db        368 CGAGAAGGCAAATCCCACTTTCAGATCCTGCTGGACGTGGTCCAGTTCCTCCCTGAAGAC 427

Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              :::   ::::::      :::   ::::::   |||      :::|||      |||
Db        428 ATCATCATTCAGACCTTCGAAGGCTGGCTACTGATAAAAGCACAACACGGAACCAGAATG 487

Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
              |||:::||||||:::|||||||||   |||   |||:::|||:::||||||      |||
Db        488 GATGAGCACGGTTTTATCTCAAGAAGCTTCACCCGACAGTACAAACTACCAGATGGCGTG 547

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
              :::       ||||||      |||      ||||||:::|||
Db        548 GAAATCAAAGATTTGTCTGCAGTCCTCTGTCATGATGGAATTTTG 592

RESULT 28
US-09-949-016-14365
; Sequence 14365, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14365
;   LENGTH: 4763
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14365

Alignment Scores:
Pred. No.:              3.47e-11      Length:         4763
Score:                  158.00        Matches:        31
Percent Similarity:     60.00%        Conservative:   14
Best Local Similarity:  41.33%        Mismatches:     30
Query Match:            17.25%        Indels:         0
DB:                     4             Gaps:           0

US-10-657-740-1 (1-173) x US-09-949-016-14365 (1-4763)

Qy         65 ArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
              |||       :::       |||   |||   |||||||||       |||   |||||||||
Db       2368 CGAGAAGGCAAATCCCACTTTCAGATCCTGCTGGACGTGGTCCAGTTCCTCCCTGAAGAC 2427

Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              :::   ::::::      :::   ::::::   |||      :::|||      |||
Db       2428 ATCATCATTCAGACCTTCGAAGGCTGGCTACTGATAAAAGCACAACACGGAACCAGAATG 2487
```

```
Qy        105 AspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
              |||:::||||||:::|||||||||    |||   |||:::|||:::||||||      |||
Db       2488 GATGAGCACGGTTTTATCTCAAGAAGCTTCACCCGACAGTACAAACTACCAGATGGCGTG 2547

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
              :::       ||||||      |||       ||||||:::|||
Db       2548 GAAATCAAAGATTTGTCTGCAGTCCTCTGTCATGATGGAATTTTG 2592

RESULT 29
US-09-513-999C-10502
; Sequence 10502, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 10502
;   LENGTH: 333
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-10502

Alignment Scores:
Pred. No.:              7.69e-12        Length:         333
Score:                  149.50          Matches:        34
Percent Similarity:     60.56%          Conservative:   9
Best Local Similarity:  47.89%          Mismatches:     19
Query Match:            16.32%          Indels:         10
DB:                     4               Gaps:           1

US-10-657-740-1 (1-173) x US-09-513-999C-10502 (1-333)

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              |||||||||   |||||||||||:::|||||||||||||||||   |||   |||:::
Db         27 GGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTCACGCGGAAA 86

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              |||   ||||||      ||||||   :::   :::|||   |||||||||   :::|||
Db         87 TACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCCCCTGAGGGC 146

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              ||||||      ||||||                          ||||||:::
Db        147 ACACTGACCGTGGA-GGCCCC-------------------------CATGCCCAAGCT 178

Qy        158 AlaIleProValSerArgGluGluLysProThr 168
              :::      |||   |||:::      |||:::
Db        179 AGCCACGCAGTCCAACGAGATCACCATCCCAGT 211

RESULT 30
US-09-949-016-16734
; Sequence 16734, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16734
;   LENGTH: 5688
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16734

Alignment Scores:
Pred. No.:              1.72e-09        Length:         5688
Score:                  147.00          Matches:        42
Percent Similarity:     45.60%          Conservative:   15
Best Local Similarity:  33.60%          Mismatches:     24
Query Match:            16.05%          Indels:         44
DB:                     4               Gaps:           5

US-10-657-740-1 (1-173) x US-09-949-016-16734 (1-5688)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
              |||           ||||||                    :::||||||   ||||||
Db       2129 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 2188

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
              ||||||||||||   ||||||                    ||||||   |||
Db       2189 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 2224

Qy         40 ------------------------------------------------------------ 40

Db       2225 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 2284

Qy         41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                                   :::    :::|||   |||   :::::::|||   |||
Db       2285 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 2344

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
              |||   ||||||:::||||||:::|||          |||::::::   :::
Db       2345 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 2401

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
              |||||||||   ||||||:::|||::::::||||||||||||   :::|||      |||
Db       2402 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 2461

Qy         95 GluIleHisGlyLys 99
              ||||||   |||:::
Db       2462 GAGATCACCGGTGAG 2476

RESULT 31
US-09-949-016-89647/c
; Sequence 89647, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 89647
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-89647
```

```
Alignment Scores:
Pred. No.:              2.35e-10        Length:         601
Score:                  142.00          Matches:        25
Percent Similarity:     80.49%          Conservative:   8
Best Local Similarity:  60.98%          Mismatches:     8
Query Match:            15.50%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x US-09-949-016-89647 (1-601)

Qy          64 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 83
               :::|||   ::::::|||:::|||   :::   ||||||||||||||||||||||||
Db         427 ATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTGGATGTGAAGCACTTCTCCCCAGAG 368

Qy          84 AspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArg 103
               :::|||   |||||||||      |||   :::|||:::|||||||||||||   ||||||
Db         367 GAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAGGTGCATGGAAAACATGAAGAGCGC 308

Qy         104 Gln 104
               |||
Db         307 CAG 305

RESULT 32
US-09-270-767-563
; Sequence 563, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 563
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-563

Alignment Scores:
Pred. No.:              3.84e-10        Length:         599
Score:                  140.50          Matches:        38
Percent Similarity:     48.44%          Conservative:   24
Best Local Similarity:  29.69%          Mismatches:     53
Query Match:            15.34%          Indels:         13
DB:                     4               Gaps:           3

US-10-657-740-1 (1-173) x US-09-270-767-563 (1-599)

Qy          18 TyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeu 37
               |||      ||||||:::      :::            |||   ::::::||||||
Db         167 TACCGTCGTCGGCTGTGGCGCGATTGG---------GATTTGAACGACTGGGACTTG--- 214

Qy          38 ProPheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeu 57
                           ||||||::::::::::||||||
Db         215 ----------------------CCCTACTGGAAGCGATCACTATCTCACGTTGGATCC 250

Qy          58 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
               :::|||   |||         :::|||   |||         :::||||||
Db         251 GCACCCGATCTAAGTCGGGTAATTGTCGGAAAGGATGGTTTCGAGGCCAATGTGGATGTG 310

Qy          78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
                  |||   |||   :::::::::||||||          |||   |||   :::
Db         311 CACCTGTTCAAGCCCTATGAGATTAGCGTGAAGACCTCAGGCGACACTGTGGTCGTGGAG 370

Qy          98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
               ||||||   :::|||:::|||         ::::::   |||         :::|||
Db         371 GCCAAGCACGAGAAGCGACGTGATGGTGACACCTTCGTGGGTCGCCACATCGTCAAGCGG 430

Qy         118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
               :::   ||||||               :::   :::         ||||||:::||||||
Db         431 TTCGTCCTGCCCCGAGGATACTATCCCAACGATGTGCGATCGGAACTGTCGTCCGATGGC 490

Qy         138 MetLeuThrPhe---CysGlyPro 144
               :::||||||      |||   |||
Db         491 ATACTTACCGTCAAGTGTCCGCCG 514

RESULT 33
US-09-270-767-15845
; Sequence 15845, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15845
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-15845

Alignment Scores:
Pred. No.:              3.84e-10        Length:         599
Score:                  140.50          Matches:        38
Percent Similarity:     48.44%          Conservative:   24
Best Local Similarity:  29.69%          Mismatches:     53
Query Match:            15.34%          Indels:         13
DB:                     4               Gaps:           3

US-10-657-740-1 (1-173) x US-09-270-767-15845 (1-599)

Qy          18 TyrProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeu 37
               |||      ||||||:::      :::            |||   ::::::||||||
Db         167 TACCGTCGTCGGCTGTGGCGCGATTGG---------GATTTGAACGACTGGGACTTG--- 214

Qy          38 ProPheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeu 57
                           ||||||::::::::::||||||
Db         215 ----------------------CCCTACTGGAAGCGATCACTATCTCACGTTGGATCC 250

Qy          58 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
               :::|||   |||         :::|||   |||         :::||||||
Db         251 GCACCCGATCTAAGTCGGGTAATTGTCGGAAAGGATGGTTTCGAGGCCAATGTGGATGTG 310

Qy          78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
                  |||   |||   :::::::::||||||          |||   |||   :::
Db         311 CACCTGTTCAAGCCCTATGAGATTAGCGTGAAGACCTCAGGCGACACTGTGGTCGTGGAG 370

Qy          98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
               ||||||   :::|||:::|||         ::::::   |||         :::|||
Db         371 GCCAAGCACGAGAAGCGACGTGATGGTGACACCTTCGTGGGTCGCCACATCGTCAAGCGG 430

Qy         118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
               :::   ||||||               :::   :::         ||||||:::||||||
Db         431 TTCGTCCTGCCCCGAGGATACTATCCCAACGATGTGCGATCGGAACTGTCGTCCGATGGC 490

Qy         138 MetLeuThrPhe---CysGlyPro 144
               :::||||||      |||   |||
Db         491 ATACTTACCGTCAAGTGTCCGCCG 514

RESULT 34
US-09-513-999C-1052
; Sequence 1052, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
```

```
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1052
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 301..564
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 203
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 204
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 330
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 372
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 390
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 394
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 10
;   OTHER INFORMATION: Xaa=Phe or Leu
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 32
;   OTHER INFORMATION: Xaa=Leu or Val
US-09-513-999C-1052

Alignment Scores:
Pred. No.:                6.76e-09        Length:         566
Score:                    131.50          Matches:        33
Percent Similarity:       63.77%          Conservative:   11
Best Local Similarity:    47.83%          Mismatches:     14
Query Match:              14.36%          Indels:         11
DB:                       4               Gaps:           5

US-10-657-740-1 (1-173) x US-09-513-999C-1052 (1-566)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   ||||||:::                 ||||||:::   |||
Db        301 ATGGACATCGCCATCCACCACCCCTTC--------------TTKCCTTTCCACTCCCCC 345

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   |||       |||
Db        346 AGCCGCCTCTTTGACCAGTTCTTCGGRGAGCACCTGTTGGAGTCKGATSTTTTCCCG--- 402

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||::::::::::||||||:::|||   :::       |||       |||       :::
Db        403 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 462

Qy         56 ValLeuAspSerGlyIleSerGluVal 64
                    |||:::|||:::|||||||||
Db        463 TGGTTTGACACTGGACTCTCAGAGGTG 489

RESULT 35
```

```
US-09-669-751-69
; Sequence 69, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
;  APPLICANT:  Greenspan, Ralph J.
;  TITLE OF INVENTION:  Methods for Identifying Compounds for
;  TITLE OF INVENTION:  Motion Sickness, Vertigo and Other Disorders Related to
;  TITLE OF INVENTION:  Balance and the Perception of Gravity
;  FILE REFERENCE:  P-NI 3864
;  CURRENT APPLICATION NUMBER: US/09/669,751
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER:  US 60/168,579
;  PRIOR FILING DATE:  1999-12-02
;  NUMBER OF SEQ ID NOS:  261
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 69
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Drosophila
US-09-669-751-69

Alignment Scores:
Pred. No.:                3.92e-08        Length:         495
Score:                    125.50          Matches:        34
Percent Similarity:       46.15%          Conservative:   14
Best Local Similarity:    32.69%          Mismatches:     39
Query Match:              13.70%          Indels:         17
DB:                       4               Gaps:           3

US-10-657-740-1 (1-173) x US-09-669-751-69 (1-495)

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro--------- 38
              |||   :::       ||||||   |||:::       :::||||||
Db        179 TTGCTGGAGGATGACTTCGGTTTTGGCGTCCATGCCCACGATCTGTTCCATCCGCGTCGC 238

Qy         39 ---PheLeuSerSerThrIle--------------------SerProTyrTyrArgGln 50
                 |||   :::|||:::                      |||||||||   |||
Db        239 CTGCTACTGCCCAACACCCTGGGACTGGGTCGTCGTCGCTATTCGCCGTACGAGAGGAGC 298

Qy         51 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSer------------ 66
                              :::             |||         :::
Db        299 CATGGCCACCACAATCAAATGTCACGTCGCGCGTCGGGAGGTCCAAACGCTCTGCTGCCC 358

Qy         67 ------AspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAsp 84
                   :::|||   |||   :::   :::||||||       |||   |||   :::
Db        359 GCCGTGGGCAAAGATGGCTTCCAGGTGTGCATGGATGTGTCGCAGTTCAAGCCCAACGAG 418

Qy         85 LeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGln 104
              ||||||||||||||   |||:::   |||   :::   |||||||   ||||||:::
Db        419 CTGACCGTCAAGGTGGTGGACAACACCGTGGTGGTAGAGGGCAAGCACGAGGAGCGCGAG 478

Qy        105 AspAspHisGly 108
              |||   ||||||
Db        479 GACGGCCATGGA 490

RESULT 36
US-09-949-016-89648/c
; Sequence 89648, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 89648
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-89648

Alignment Scores:
Pred. No.:              6.32e-07        Length:         601
Score:                  118.00          Matches:        21
Percent Similarity:     82.76%          Conservative:   3
Best Local Similarity:  72.41%          Mismatches:     5
Query Match:            12.88%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x US-09-949-016-89648 (1-601)

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        601 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 542

Qy         96 IleHisGlyLysHisAsnGluArgGln 104
              :::|||||||||||||   |||||||||
Db        541 GTGCATGGAAAACATGAAGAGCGCCAG 515

RESULT 37
US-09-949-016-17284/c
; Sequence 17284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17284
;   LENGTH: 12214
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17284

Alignment Scores:
Pred. No.:              0.000221        Length:         12214
Score:                  115.00          Matches:        35
Percent Similarity:     51.35%          Conservative:   22
Best Local Similarity:  31.53%          Mismatches:     36
Query Match:            12.55%          Indels:         18
DB:                     4               Gaps:           5

US-10-657-740-1 (1-173) x US-09-949-016-17284 (1-12214)

Qy         54 ArgThrValLeuAspSerGlyIleSerGluValArgSerAsp------ArgAspLysPhe 71
              |||   :::   ||||||      ::::::      :::|||      ||||||   |||
Db        376 CGGCTGCTCAGGGACAGTCCAGCGGCTCAGGAGGACAATGACCATGCCAGAGACGGTTTC 317

Qy         72 ValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAsp 91
              :::   ||||||        |||:::||||||:::|||   |||:::|||
Db        316 CAAATGAAGCTGGATGCCCACGGCTTCGCCCCGGAGGAACTGGTGGTGCAGGTGGATGGC 257

Qy         92 AspPheValGluIleHisGlyLysHisAsn-----------------GluArgGlnAsp 105
              ::::::   :::   |||:::                        ||||||
Db        256 CAATGGCTGATGGTGACCGGACAGCAGCAACTGGACGTCAGGGACCCGGAAAGGGTCAGT 197

Qy        106 AspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAsp 125
                      :::|||::::::   ||||||:::      ||||||||||||:::
Db        196 TACCGC------ATGTCACAGAAGGTGCACCGGAAA---ATGCTACCGTCCAACCTGAGT 146

Qy        126 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly------ 143
              :::|||::::::|||   |||:::      |||   |||          |||
Db        145 CCTACCGCCATGACCTGCTGCCTGACCCCCTCCGGGCAGCTGTGGGTCAGAGGCCAGTGT 86

Qy        144 --------------ProLysIleGlnThrGly 149
                           |||:::   |||||||||
Db         85 GTGGCGCTGGCCCTCCCTGAAGCCCAAACAGGA 53

RESULT 38
US-09-620-312D-669
; Sequence 669, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 669
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (81)..(578)
US-09-620-312D-669

Alignment Scores:
Pred. No.:              9.95e-06        Length:         719
Score:                  110.50          Matches:        23
Percent Similarity:     54.32%          Conservative:   21
Best Local Similarity:  28.40%          Mismatches:     34
Query Match:            12.06%          Indels:         3
DB:                     4               Gaps:           1

US-10-657-740-1 (1-173) x US-09-620-312D-669 (1-719)

Qy         60 GlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHis 79
              |||         :::::::::      |||   :::         :::|||||:::
Db        276 GGGGCAGGCAACATCAAGACCCTAGGAGACGCCTATGAGTTTGCGGTGGACGTGAGAGAC 335

Qy         80 PheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLys 99
              ||||||||||||||:::   |||         ::::::   :::|||:::
Db        336 TTCTCACCTGAAGACATCATTGTCACCACCTCCAACAACCACATCGAGGTGCGGGCT--- 392

Qy        100 HisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArg 119
                     |||:::            |||   :::         |||      :::   :::
```

```
Db        393 ------GAGAAGCTGGCGGCTGACGGCACTGTCATGAACACCTTCGCTCACAAGTGCCAG 446

Qy        120 LeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
              ||||||   :::||||||   ::::::::::   :::|||      ||||||   |||
Db        447 CTGCCGGAGGACGTGGACCCGACGTCGGTGACCTCGGCTCTGCGGGAGGACGGCAGCCTC 506

Qy        140 Thr 140
              |||
Db        507 ACT 509

RESULT 39
US-09-949-016-17282/c
; Sequence 17282, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17282
;   LENGTH: 9615
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17282

Alignment Scores:
Pred. No.:              0.000778        Length:         9615
Score:                  110.00          Matches:        31
Percent Similarity:     51.58%          Conservative:   18
Best Local Similarity:  32.63%          Mismatches:     30
Query Match:            12.01%          Indels:         16
DB:                     4               Gaps:           4

US-10-657-740-1 (1-173) x US-09-949-016-17282 (1-9615)

Qy         68 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 87
              ||||||   |||   :::   ||||||        |||:::||||||:::|||   |||
Db       9588 AGAGACGGTTTCCAAATGAAGCTGGATGCCCACGGCTTCGCCCCGGAGGAACTGGTGGTG 9529

Qy         88 LysValGlnAspAspPheValGluIleHisGlyLysHisAsn---------------- 101
              :::|||         ::::::   :::   |||:::
Db       9528 CAGGTGGATGGCCAATGGCTGATGGTGACCGGACAGCAGCAACTGGACGTCAGGGACCCG 9469

Qy        102 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
              ||||||                ::::|||::::::   ||||||:::      ||||||
Db       9468 GAAAGGGTCAGTTACCGC------ATGTCACAGAAGGTGCACCGGAAA---ATGCTACCG 9418

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
              ||||||:::      :::|||::::::|||   |||:::      |||   |||
Db       9417 TCCAACCTGAGTCCTACCGCCATGACCTGCTGCCTGACCCCCTCCGGGCAGCTGTGGGTC 9358

Qy        142 CysGly--------------------ProLysIleGlnThrGly 149
                 |||                    |||:::   |||||||||
Db       9357 AGAGGCCAGTGTGTGGCGCTGGCCCTCCCTGAAGCCCAAACAGGA 9313

RESULT 40
US-09-949-016-14287/c
; Sequence 14287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14287
;   LENGTH: 7125
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(7125)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14287

Alignment Scores:
Pred. No.:              0.00129         Length:         7125
Score:                  107.00          Matches:        18
Percent Similarity:     69.05%          Conservative:   11
Best Local Similarity:  42.86%          Mismatches:     13
Query Match:            11.68%          Indels:         0
DB:                     4               Gaps:           0

US-10-657-740-1 (1-173) x US-09-949-016-14287 (1-7125)

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
              :::|||   ||||||:::|||            ||||||      ||||||||||||
Db       130 GCAGGGGCCTCCGAGCTTAGGCTCAGTGAGGGCAAGTTCCAGGCATTTCTGGATGTGAGC 71

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||:::|||:::::::::||||||:::      |||:::   :::|||:::
Db         70 CACTTTACCCCAGACGAGGTGACTGTGAGGACTGTGGATAACCTGCTGGAGGTGTCTGCC 11

Qy         99 LysHis 100
              :::|||
Db         10 CGGCAC 5

RESULT 41
US-09-949-016-175645
; Sequence 175645, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 175645
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-175645

Alignment Scores:
Pred. No.:              0.000632        Length:         601
Score:                  97.00           Matches:        29
Percent Similarity:     39.33%          Conservative:   6
Best Local Similarity:  32.58%          Mismatches:     13
```

```
Query Match:            10.59%      Indels:        41
DB:                     4           Gaps:          2

US-10-657-740-1 (1-173) x US-09-949-016-175645 (1-601)

Qy         90 GlnAspAspPhe---ValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGly 108
              |||   ||||||   :::         |||||||||   |||||||||||:::||||||
Db        333 CAGTCTGATTTCCCTCTTCCCCCCAAAGGCAAGCACGAGGAGCGGCAGGACGAGCATGGC 392

Qy        109 TyrIleSerArgGluPheHisArgArgTyr---------------------------- 118
              ||||||||||||   |||   |||:::|||
Db        393 TACATCTCCCGGTGCTTCACGCGGAAATACACGTGAGTCCTGGCGCCAGGTCGGGGTGGG 452

Qy        118 ------------------------------------------------------------ 118

Db        453 TGGGTGGCGTGGGGGTGGGGTCAGGGAAGAGGGCACAGGGACCCACCCGGTGTGTAATGT 512

Qy        119 ----------------------------ArgLeuProSerAsnValAspGlnSerAlaLe 129
                                           |||||||||      ||||||   :::   ::
Db        513 AACGCTTGCCTTTCCTCTCTGCACGTCCAGGCTGCCCCCCGGTGTGGACCCCACCCAAGT 572

Qy        129 uSerCysSerLeuSerAlaAspGly 137
              :|||   |||||||||   :::|||
Db        573 TTCCTCCTCCCTGTCCCCTGAGGGC 597

RESULT 42
US-09-107-532A-670
; Sequence 670, Application US/09107532A
; Patent No. 6583275
;  GENERAL INFORMATION:
;       APPLICANT: Lynn A Doucette-Stamm  and David Bush
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                           ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 7310
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;            STREET: 100 Beaver Street
;            CITY: Waltham
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02354
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: PC
;            OPERATING SYSTEM: <Unknown>
;            SOFTWARE: ASCII
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/107,532A
;            FILING DATE: 30-Jun-1998
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/085,598
;            FILING DATE: 14 May 1998
;            APPLICATION NUMBER: 60/051571
;            FILING DATE: July 2, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ariniello, Pamela Deneke
;            REGISTRATION NUMBER: 40,489
;            REFERENCE/DOCKET NUMBER: GTC-012
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (781)893-5007
;            TELEFAX: (781)893-8277
;  INFORMATION FOR SEQ ID NO: 670:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 420 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: circular
;       MOLECULE TYPE: DNA (genomic)
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: Enterococcus faecium
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION: (B) LOCATION 1...420
;       SEQUENCE DESCRIPTION: SEQ ID NO: 670:
US-09-107-532A-670

Alignment Scores:
Pred. No.:              0.000492    Length:        420
Score:                  96.00       Matches:       31
Percent Similarity:     45.16%      Conservative:  25
Best Local Similarity:  25.00%      Mismatches:    46
Query Match:            10.48%      Indels:        22
DB:                     4           Gaps:          6

US-10-657-740-1 (1-173) x US-09-107-532A-670 (1-420)

Qy         35 AspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
              |||:::   |||      :::         ||||||   :::         |||
Db         31 GACATGTTCCCTGATTTCAATGACGTGTTTAGTCCAGCATTCAATGATTTTTTA------ 84

Qy         55 ThrValLeuAspSerGlyIleSer--------------GluValArgSerAspArgAsp 69
                        |||:::|||                 ::::::      :::
Db         85 --------------GGTGTTTCAAGTTATCCAAAAGTAGATTTAGTCGAAAATGAAAAA 129

Qy         70 LysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysVal 89
              ::::::   :::      |||:::                |||||   ||||||:::
Db        130 GAATACAAACTAACTGCTGATATGCCTGGCTGTGACAAAGAAGACACGACAGTTGAATAT 189

Qy         90 GlnAspAspPheValGluIle---------HisGlyLysHisAsnGluArgGlnAspAsp 106
                 |||:::   :::   |||         |||      |||   |||   ::::::|||
Db        190 TCAGATAATACTTTGACGATCTCCGCCAATCATGAGTCCCATACAGAAGATAAAGAAGAT 249

Qy        107 HisGlyTyrIleSerArgGluPheHis--------------ArgArgTyrArgLeuPro 121
                   |||:::   :::|||   |||                |||   :::   ||||||
Db        250 GGCAACTATGTACGGAAAGAAAGACATTCGGTCTCTTACAAACGTTCATTCTATCTTCCT 309

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 141
              |||||||||:::      ::::::   :::         :::|||:::|||
Db        310 ---AATGTTGATGAAGAAAAAATCACTGGAACCTTTAAA---AACGGTGTGCTGAAATTG 363

Qy        142 CysGlyProLys 145
                   ||||||
Db        364 GTATTACCTAAA 375

RESULT 43
US-07-748-761-2
; Sequence 2, Application US/07748761
; Patent No. 5304490
;  GENERAL INFORMATION:
;    APPLICANT:  Grierson, Donald
;    APPLICANT:  Fray, Rupert G
;    APPLICANT:  Lycett, Grantley W
;    APPLICANT:  Bird, Colin R
;    APPLICANT:  Ray, John A
;    APPLICANT:  Schuch, Wolfgang W
;    TITLE OF INVENTION:  DNA, DNA constructs, cells and plants
;    TITLE OF INVENTION:  derived therefrom
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman, Darby and Cushman
;      STREET:  Eleventh floor, 1615 L Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/748,761
;      FILING DATE:  19910823
;      CLASSIFICATION:  800
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9018612.3
;      FILING DATE:  24-AUG-1990
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  652 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-07-748-761-2

Alignment Scores:
Pred. No.:              0.00228       Length:        652
Score:                  93.50         Matches:       36
Percent Similarity:     43.97%        Conservative:  26
Best Local Similarity:  25.53%        Mismatches:    56
Query Match:            10.21%        Indels:        23
DB:                     1             Gaps:          7

US-10-657-740-1 (1-173) x US-07-748-761-2 (1-652)

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              |||:::|||                  :::|||:::           ||||||
Db         60 AGAATTTTCGGCGATCGACGAAGCAGCAGCATGTTCGAT------------CCATTTTCA 107

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
                        |||   :::|||:::     |||   :::     :::||||||
Db        108 ATTGACGTATTTGATCCA---TTCAGGGAATTAGGCTTCCCAAGTACC---AATTCAGGG 161

Qy         61 IleSerGluVal----------------------ArgSerAspArgAspLysPheVal 72
                |||                           :::       :::      |||
Db        162 GAGAGCTCTGCATTTGCCAACACACGAATAGACTGGAAGGAAACTCCAGAACCTCATGTG 221

Qy         73 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 92
                  :::|||:::               |||::::::   |||:::|||::::::|||
Db        222 TTCAAGGTTGATCTTCCAGGGCTTAAGAAGGAGGAAGTCAAAGTGGAAGTCGAGGAGGAT 281

Qy         93 PheVal---GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGly--------- 108
                |||   :::|||   |||:::   |||     ::::::|||
Db        282 AGGGTTCTTCAGATCAGCGGAGAGAGGAACGTGGAGAAGGAAGATAAGAATGATAAGTGG 341

Qy        109 ------------TyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnVal 124
                        |||   :::|||   ||||||:::||||||||   |||
Db        342 CATCGCATGGAGCGAAGCAGCGGGAAATTCATGAGGAGATTTAGACTTCCGGAGAATGCA 401

Qy        125 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 144
                      :::     |||:::     :::|||:::||||||           |||
Db        402 AAGATGGATCAAGTTAAGGCGTCTATGGAG---AATGGAGTGCTTACTGTTACTGTTCCA 458

Qy        145 Lys 145
              |||
Db        459 AAG 461

RESULT 44
US-09-252-991A-12236
; Sequence 12236, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 12236
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12236

Alignment Scores:
Pred. No.:              0.00216       Length:        465
Score:                  92.00         Matches:       34
Percent Similarity:     40.58%        Conservative:  22
Best Local Similarity:  24.64%        Mismatches:    50
Query Match:            10.04%        Indels:        32
DB:                     4             Gaps:          5

US-10-657-740-1 (1-173) x US-09-252-991A-12236 (1-465)

Qy          8 ProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePhe 27
              |||   |||:::   :::::::|||                  ||||||:::|||
Db         40 CCGCTGTTCCGTCATTCCGTAGGC--------------------TTCGATCGCTTC--- 75

Qy         28 GlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyr 47
                       |||||||||     |||              ||||||     ||||||
Db         76 ---AACGATCTGTTCGAGTCGGCCCTGCGCAATGAGGCCGGGAGTACCTACCCGCCCTAC 132

Qy         48 TyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAsp 67
                                                       :::      :::
Db        133 ----------------------------------------AACGTCGAAAAGCACGGT 150

Qy         68 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 87
                 |||::::::   |||   :::              |||      |||||||||   :::
Db        151 GACGACGAGTATCGCATCGTCATCGCCGCCGCCGGCTTCCAGGAAGAAGACCTGGACCTG 210

Qy         88 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 107
                 :::|||:::          :::   :::   |||          |||:::      |||:::
Db        211 CAGGTCGAGCGCGGCGTGCTGACCGTCAGTGGCGGCAAGCGCGAGAAGTCCACGGACAAC 270

Qy        108 ------------------GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
                             |||          |||   |||          :::||||||
Db        271 GTGACCTACCTGCACCAGGGCATCGCCCAGCGCGCCTTCAAGCTGTCGTTCCGCCTCGCC 330

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
                  :::::::::       |||   |||              |||:::|||:::|||
Db        331 GATCATATCGAGGTCAAGGCGGCTTCGCTG---------GCCAACGGCCTGCTG 375

RESULT 45
US-09-252-991A-12373/c
; Sequence 12373, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 12373
;   LENGTH: 522
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12373

Alignment Scores:
Pred. No.:              0.00261       Length:        522
Score:                  92.00         Matches:       34
Percent Similarity:     40.58%        Conservative:  22
Best Local Similarity:  24.64%        Mismatches:    50
Query Match:            10.04%        Indels:        32
```

```
DB:                    4              Gaps:          5

US-10-657-740-1 (1-173) x US-09-252-991A-12373 (1-522)

Qy          8 ProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhePhe 27
              |||  |||:::  ::::::|||                 ||||||:::|||
Db        429 CCGCTGTTCCGTCATTCCGTAGGC--------------------TTCGATCGCTTC--- 394

Qy         28 GlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSerProTyr 47
                |||||||||      |||                 ||||||      ||||||
Db        393 ---AACGATCTGTTCGAGTCGGCCCTGCGCAATGAGGCCGGGAGTACCTACCCGCCCTAC 337

Qy         48 TyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAsp 67
                                                    :::      :::
Db        336 ----------------------------------------AACGTCGAAAAGCACGGT 319

Qy         68 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 87
                ||:::::   |||   :::          |||      |||||||||   :::
Db        318 GACGACGAGTATCGCATCGTCATCGCCGCCGCCGGCTTCCAGGAAGAAGACCTGGACCTG 259

Qy         88 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 107
              :::|||:::         :::   :::   |||          |||:::      |||:::
Db        258 CAGGTCGAGCGCGGCGTGCTGACCGTCAGTGGCGGCAAGCGCGAGAAGTCCACGGACAAC 199

Qy        108 ------------------GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 121
                               |||          |||   |||          :::|||||||
Db        198 GTGACCTACCTGCACCAGGGCATCGCCCAGCGCGCCTTCAAGCTGTCGTTCCGCCTCGCC 139

Qy        122 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeu 139
              :::::::::      |||   |||          |||:::|||:::|||
Db        138 GATCATATCGAGGTCAAGGCGGCTTCGCTG---------GCCAACGGCCTGCTG 94

RESULT 46
US-09-489-039A-3463
; Sequence 3463, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 3463
;   LENGTH: 576
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3463

Alignment Scores:
Pred. No.:              0.00306       Length:         576
Score:                  92.00         Matches:        30
Percent Similarity:     38.41%        Conservative:   28
Best Local Similarity:  19.87%        Mismatches:     59
Query Match:            10.04%        Indels:         34
DB:                     4             Gaps:           4

US-10-657-740-1 (1-173) x US-09-489-039A-3463 (1-576)

Qy         35 AspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeu------ 52
              :::|||             :::      :::||||||   ||||||   ::::::
Db        151 GAGCTATTTACTATGCGTAACTTCGATCTTTCCCCGCTATATCGTTCAGCCATTGGTTTC 210

Qy         53 ------PheArgThrValLeuAspSerGlyIleSerGluValArgSer------------ 66
                     :::|||::::::      |||:::
Db        211 GACCGCCTGTTCAACCTACTGGAAAACAATCAAAGCCAGAGCAACGGCGGCTACCCTCCG 270

Qy         67 -----------------AspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
                              |||   :::   :::   |||   :::   |||       |||
Db        271 TATAACGTCGAGCTGGTAGACGAAAACCACTATCGCATCGCTATCGCGGTGGCTGGCTTT 330

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              :::     :::|||   :::          ||||||:::   :::   :::   |||   |||
Db        331 GCTGAAAGCGAGCTGGAGATCACTGCCCAGGACAATCTGCTGATCGTCAAAGGCGCCCAC 390

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIle--------------SerArgGluPheHis 115
                       |||   :::      |||:::                  |||   |||
Db        391 GCCGCCGAGCAGAAAGAGCGGACCTACCTGTATCAGGGGATCGCCGAGCGCAACTTCGAG 450

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::::::::|||       |||:::
Db        451 CGCAAATTCCAGCTGGCGGAAAACATTCACGTCCGC----------------------- 486

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
                                  |||      :::      ||||||
Db        487 --------------------GGCGCCAATCTGGTGAACGGCCTGCTGTATATCGATCTG 525

Qy        156 GluArgAlaIleProValSerArgGluGluLys 166
              ||||||   ||||||   :::   :::   :::
Db        526 GAACGGGTGATCCCGGAAGCGAACAAGCCGCGC 558

RESULT 47
US-09-949-016-13807/c
; Sequence 13807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13807
;  LENGTH: 168971
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(168971)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13807

Alignment Scores:
Pred. No.:              112           Length:         168971
Score:                  88.00         Matches:        30
Percent Similarity:     46.99%        Conservative:   9
Best Local Similarity:  36.14%        Mismatches:     30
Query Match:            9.61%         Indels:         14
DB:                     4             Gaps:           4

US-10-657-740-1 (1-173) x US-09-949-016-13807 (1-168971)

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg--- 117
              |||      |||  ::::::      |||          |||      |||||||||
Db     167465 AAACCTACTGAGGGTGAGGAGGCACCGGGA---GCCGGGAGGAGAGGGCATCGCAGGAGA 167409

Qy        118 ---TyrArgLeuProSerAsnValAspGlnSerAlaLeu--------------SerCys 131
                 :::|||||||||          |||:::||||||                :::|||
Db     167408 AAGTGGCGGCTCCCCTGCCCTCATTCCCAGGCAGCACTGGGCAGTGGTCGCGTCAACTGT 167349

Qy        132 SerLeuSer-----------------AlaAspGlyMetLeuThrPheCysGlyProLys 145
              :::||||||                        :::||||||          ||||||
Db     167348 GCTCTGAGCTGGGGCTCCAGTGACGGCACTGAGGGGATGGGAAGGTCCTGTGGGGGACTC 167289

Qy        146 IleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGlu 165
```

```
          :::  |||||||||  |||  ||||||  :::|||
Db     167288 CTAACCACCGGCCTACCAGCGTGTGTGGCCGAGTCCCTCCTTCCCTCTCCTGCCAGAATG 167229

Qy        166 LysProThr 168
              |||||||||
Db     167228 AAGCCCACC 167220

RESULT 48
US-09-902-540-6436
; Sequence 6436, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6436
;   LENGTH: 489
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-6436

Alignment Scores:
Pred. No.:              0.0169          Length:         489
Score:                  86.00           Matches:        33
Percent Similarity:     43.36%          Conservative:   16
Best Local Similarity:  29.20%          Mismatches:     39
Query Match:            9.39%           Indels:         26
DB:                     4               Gaps:           4

US-10-657-740-1 (1-173) x US-09-902-540-6436 (1-489)

Qy         63 GluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerPro 82
              ||||||:::      :::|||   |||:::       ||||||
Db        154 GAGGTGAAGGAGACCCAGGATGCCTTCATCTTCAAGGCCGATGTCCCCGGCGTGGAGGAG 213

Qy         83 GluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGlu 102
              :::||||||   :::   :::   ::::::   :::   |||   ||||||      |||
Db        214 AAGGACCTGGAAATCACCCTCGCGGAGAACCGGCTGACCATCAGCGGCAAGCGGGAAGAA 273

Qy        103 ArgGlnAspAspHisGly------TyrIleSerArgGlu--------------PheHis 115
                 :::  |||  |||      |||  :::  |||              |||
Db        274 GAGCGCCGCGACGAAGGCGACCGCTACTACGCCTACGAGCGCAACTATGGCTCGTTCAGC 333

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||   :::   ||||||      |||                          :::|||
Db        334 CGCACCTTCACGCTCCCCCGGGGTGTG-------------------------AATGCC 366

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              |||                          :::|||   ||||||
Db        367 GAC----------------------AACGTCCAGGC-GGACTTCAAGAGTGGCGTCCT 401

Qy        156 GluArgAla-IleProValSerArgGluGluLysPro 167
              :::|||||| ||||||        ||||||:::|||
Db        402 CAACGTGCGCATCCCCAAGAAGTCCGAGGAGCAGCCC 438

RESULT 49
US-09-902-540-444
; Sequence 444, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 444
;   LENGTH: 3626
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(3626)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-444

Alignment Scores:
Pred. No.:              0.431           Length:         3626
Score:                  86.00           Matches:        33
Percent Similarity:     43.36%          Conservative:   16
Best Local Similarity:  29.20%          Mismatches:     39
Query Match:            9.39%           Indels:         26
DB:                     4               Gaps:           4

US-10-657-740-1 (1-173) x US-09-902-540-444 (1-3626)

Qy         63 GluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerPro 82
              ||||||:::      :::|||   |||:::       ||||||
Db       1509 GAGGTGAAGGAGACCCAGGATGCCTTCATCTTCAAGGCCGATGTCCCCGGCGTGGAGGAG 1568

Qy         83 GluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGlu 102
              :::||||||   :::   :::   ::::::   :::   |||   ||||||      |||
Db       1569 AAGGACCTGGAAATCACCCTCGCGGAGAACCGGCTGACCATCAGCGGCAAGCGGGAAGAA 1628

Qy.        103 ArgGlnAspAspHisGly------TyrIleSerArgGlu--------------PheHis 115
                 :::  |||  |||      |||  :::  |||              |||
Db       1629 GAGCGCCGCGACGAAGGCGACCGCTACTACGCCTACGAGCGCAACTATGGCTCGTTCAGC 1688

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||   :::   ||||||      |||                          :::|||
Db       1689 CGCACCTTCACGCTCCCCCGGGGTGTG-------------------------AATGCC 1721

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               |||                          :::|||   ||||||
Db       1722 GAC----------------------AACGTCCAGGC-GGACTTCAAGAGTGGCGTCCT 1756

Qy         156 GluArgAla-IleProValSerArgGluGluLysPro 167
               :::|||||| ||||||        ||||||:::|||
Db       1757 CAACGTGCGCATCCCCAAGAAGTCCGAGGAGCAGCCC 1793

RESULT 50
US-09-949-016-11770/c
; Sequence 11770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11770
;   LENGTH: 49487
;   TYPE: DNA
;   ORGANISM: Human
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(49487)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770

Alignment Scores:
Pred. No.:              57              Length:         49487
Score:                  84.00           Matches:        46
Percent Similarity:     40.74%          Conservative:   31
Best Local Similarity:  24.34%          Mismatches:     54
Query Match:            9.17%           Indels:         58
DB:                     4               Gaps:           12

US-10-657-740-1 (1-173) x US-09-949-016-11770 (1-49487)

Qy         34 TyrAspLeuLeuProPheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPhe 53
              |||    :::   |||::::::   |||   :::|||         ||||||   |||
Db       2693 TACTACATATGGCCATGGGTAAAATCA---CTCTCTCTGATTCGTAGGCAGTTTCTTCAC 2637

Qy         54 ArgThrValLeuAspSerGlyIleSerGluValArgSerAsp----------------- 67
              |||                  :::||||||:::|||   :::
Db       2636 CGC-----------------TTAAGTGAGCTGAGGTGTGAACGACCAGATAGATAATCA 2595

Qy         68 -----------------ArgAspLysPheValIlePheLeuAsp-------------- 76
                               |||::::::   :::|||::::::
Db       2594 TCCAGATCGGAAAACCCCCGAGAGCGCAGGATTATCTGGATAACGGAGCCGGATGGAGGA 2535

Qy         77 ------------ValLysHisPheSerProGlu---------AspLeuThrValLysVal 89
                      :::  |||   |||||||||         ||||||   :::
Db       2534 TTGTCTCCGTTCCTCCCACACTCAAGCCCCGAAAGGTCACCTGACCTAGTCTTATTCTTT 2475

Qy         90 GlnAspAspPheValGluIleHisGlyLys---HisAsn--------------------- 101
                 :::      :::   ||||||      ||||||
Db       2474 ACAGCAGAAATCGGGCAAACCCATGGCGCCGTCCACAATCTCTATAGCCGCCGACCTGGG 2415

Qy        102 ---------GluArgGlnAsp-----------------AspHisGlyTyrIleSerArg 112
                      |||::::::                 |||   |||||||||
Db       2414 TACAGGGGAAGCAGGCGAAACAACACACACTTCTACGCAGACCAGGGTTACATTAAGGCC 2355

Qy        113 GluPheHisArgArgTyrArgLeuProSerAsnVal--------------AspGlnSer 127
              :::   |||   |||          |||::::::                  |||:::
Db       2354 GACGCCCACACACGCCCACACACACCCACACACACACACACACCCCCACACCCCCAGACT 2295

Qy        128 AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln 147
              :::|||:::|||                          :::|||   |||:::::|||
Db       2294 TCGTTAACTTGTGATGATAAATTAGGGCGCAGGTGACACTACTGT---CCCGAACTTCAG 2238

Qy        148 Thr---------GlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGlu 164
                          :::      :::   |||:::|||          |||
Db       2237 CTGGAACGCTAGTCCAAGTCACTGTCCTAACTGCAGCTAGCACTGCCGACCTTGTCCGAG 2178

Qy        165 GluLysProThrSerAlaProSerSer 173
              |||   |||   ||||||||||||:::
Db       2177 GAGTTTCCT---TCCGCTCCCTCAACT 2154


Search completed: May 30, 2005, 08:53:36
Job time : 229.078 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2005, 05:14:41 ; Search time 1216.26 Seconds
(without alignments)
873.287 Million cell updates/sec

Title: US-10-657-740-1
Perfect score: 916
Sequence: 1 MDVTIQHPWFKRTLGPFYPS.....HAERAIPVSREEKPTSAPSS 173

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10657740/runat_27052005_165254_3380/app_query.fasta_1.590
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=150 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10657740_@CGN_1_1_904_@runat_27052005_165254_3380
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	871	95.1	1056	16	US-10-316-253-112	Sequence 112, App
2	849.5	92.7	1271	16	US-10-316-253-110	Sequence 110, App
3	702	76.6	531	13	US-10-105-427-1	Sequence 1, Appli
4	639	69.8	372	18	US-10-657-740-2	Sequence 2, Appli
5	493	53.8	548	15	US-10-101-510-250	Sequence 250, App
6	489	53.4	528	18	US-10-770-668-35	Sequence 35, Appl
7	489	53.4	691	9	US-09-954-456-514	Sequence 514, App
8	489	53.4	691	10	US-09-960-706-869	Sequence 869, App
9	489	53.4	691	10	US-09-873-319-566	Sequence 566, App
10	489	53.4	691	10	US-09-873-367C-764	Sequence 764, App
11	489	53.4	691	17	US-10-133-937-61	Sequence 61, Appl
12	489	53.4	691	17	US-10-172-118-686	Sequence 686, App
13	489	53.4	691	17	US-10-159-563-61	Sequence 61, Appl
14	489	53.4	691	17	US-10-342-887-686	Sequence 686, App
15	489	53.4	691	19	US-10-843-641A-764	Sequence 764, App
16	489	53.4	691	19	US-10-843-641A-3541	Sequence 3541, Ap
17	489	53.4	691	19	US-10-367-057-187	Sequence 187, App
18	485	52.9	528	17	US-10-152-319A-1574	Sequence 1574, Ap
19	485	52.9	689	19	US-10-486-706-455	Sequence 455, App
20	485	52.9	1247	9	US-09-917-800A-1419	Sequence 1419, Ap
21	485	52.9	1247	19	US-10-486-706-105	Sequence 105, App
22	418	45.6	449	9	US-09-960-352-4277	Sequence 4277, Ap
23	394	43.0	396	9	US-09-960-352-12622	Sequence 12622, A
24	367	40.1	211	16	US-10-029-386-26581	Sequence 26581, A
25	367	40.1	573	16	US-10-029-386-12881	Sequence 12881, A
26	364.5	39.8	411	14	US-10-198-846-10360	Sequence 10360, A
27	361	39.4	450	10	US-09-918-995-4949	Sequence 4949, Ap
28	347	37.9	579	16	US-10-029-386-1061	Sequence 1061, Ap
29	347	37.9	834	13	US-10-027-632-172414	Sequence 172414,
30	347	37.9	834	17	US-10-027-632-172414	Sequence 172414,
31	345.5	37.7	829	17	US-10-062-674-1924	Sequence 1924, Ap
32	340	37.1	480	14	US-10-226-956-320	Sequence 320, App
33	340	37.1	1820	18	US-10-723-860-5075	Sequence 5075, Ap
34	338	36.9	212	16	US-10-029-386-14764	Sequence 14764, A
35	338	36.9	421	10	US-09-918-995-4772	Sequence 4772, Ap
36	336	36.7	486	14	US-10-226-956-303	Sequence 303, App
37	336	36.7	1310	17	US-10-152-319A-1412	Sequence 1412, Ap
38	334	36.5	381	18	US-10-425-115-173881	Sequence 173881,
39	332.5	36.3	380	9	US-09-960-352-4143	Sequence 4143, Ap
40	311.5	34.0	405	10	US-09-918-995-4075	Sequence 4075, Ap
41	296.5	32.4	387	18	US-10-723-860-380	Sequence 380, App
42	295.5	32.3	543	17	US-10-621-901-2279	Sequence 2279, Ap
43	291	31.8	349	9	US-09-990-747-28	Sequence 28, Appl
44	289.5	31.6	764	18	US-10-605-498-91	Sequence 91, Appl
45	289.5	31.6	789	17	US-10-641-643-1300	Sequence 1300, Ap
46	289.5	31.6	847	14	US-10-153-668-283	Sequence 283, App
47	289.5	31.6	865	11	US-09-969-034-4480	Sequence 4480, Ap
48	289.5	31.6	865	17	US-10-172-118-626	Sequence 626, App
49	289.5	31.6	865	17	US-10-342-887-626	Sequence 626, App
50	289.5	31.6	1231	9	US-09-880-107-3865	Sequence 3865, Ap
51	289.5	31.6	1380	14	US-10-100-957A-169	Sequence 169, App
52	289	31.6	787	9	US-09-917-800A-1428	Sequence 1428, Ap
C 53	288.5	31.5	563	14	US-10-198-846-8376	Sequence 8376, Ap
54	288.5	31.5	787	17	US-10-191-803-228	Sequence 228, App
55	288.5	31.5	787	17	US-10-152-319A-1963	Sequence 1963, Ap
56	288	31.4	599	14	US-10-060-036-43	Sequence 43, Appl
57	287.5	31.4	634	19	US-10-764-420-43	Sequence 43, Appl
58	282.5	30.8	373	17	US-10-242-535A-26398	Sequence 26398, A
59	282.5	30.8	373	17	US-10-085-783A-26398	Sequence 26398, A
60	276	30.1	503	10	US-09-911-904-146	Sequence 146, App
C 61	266.5	29.1	604	9	US-09-917-800A-981	Sequence 981, App
62	266	29.0	410	18	US-10-468-091-35	Sequence 35, Appl
63	259.5	28.3	654	19	US-10-204-921-16	Sequence 16, Appl
64	257	28.1	1700	9	US-09-813-358-2	Sequence 2, Appli
65	257	28.1	1700	10	US-09-997-279-2	Sequence 2, Appli
66	251	27.4	549	10	US-09-991-936-341	Sequence 341, App
67	250.5	27.3	393	18	US-10-425-115-63403	Sequence 63403, A
68	247	27.0	550	17	US-10-621-901-2194	Sequence 2194, Ap
69	241.5	26.4	708	10	US-09-814-353-14770	Sequence 14770, A

	70	238	26.0	874	17	US-10-133-937-62 Sequence 62, Appl
	71	238	26.0	874	17	US-10-159-563-62 Sequence 62, Appl
	72	237	25.9	605	19	US-10-840-038-2 Sequence 2, Appli
	73	237	25.9	736	19	US-10-840-038-3 Sequence 3, Appli
	74	234.5	25.6	614	10	US-09-814-353-2040 Sequence 2040, Ap
	75	234.5	25.6	614	10	US-09-814-353-8386 Sequence 8386, Ap
	76	233.5	25.5	954	10	US-09-814-353-20273 Sequence 20273, A
	77	225.5	24.6	542	17	US-10-621-901-2146 Sequence 2146, Ap
	78	223.5	24.4	500	10	US-09-918-995-22134 Sequence 22134, A
	79	223	24.3	467	14	US-10-060-036-34 Sequence 34, Appl
	80	220.5	24.1	482	10	US-09-918-995-4433 Sequence 4433, Ap
c	81	212	23.1	49634	11	US-09-968-007A-514 Sequence 514, App
c	82	212	23.1	49634	19	US-10-843-641A-6984 Sequence 6984, Ap
	83	205.5	22.4	212	18	US-10-482-158-43 Sequence 43, Appl
c	84	198.5	21.7	396	9	US-09-954-531-845 Sequence 845, App
c	85	198.5	21.7	396	9	US-09-954-531-1265 Sequence 1265, Ap
c	86	198.5	21.7	396	19	US-10-843-641A-1912 Sequence 1912, Ap
c	87	198.5	21.7	396	19	US-10-843-641A-2332 Sequence 2332, Ap
	88	192	21.0	6806	17	US-10-205-331-105 Sequence 105, App
	89	191.5	20.9	445	10	US-09-918-995-10509 Sequence 10509, A
	90	188	20.5	4206	15	US-10-101-510-288 Sequence 288, App
	91	188	20.5	4206	18	US-10-283-975A-93 Sequence 93, Appl
	92	187	20.4	2632	13	US-10-044-090-312 Sequence 312, App
	93	182.5	19.9	487	17	US-10-369-493-30110 Sequence 30110, A
	94	182.5	19.9	487	17	US-10-369-493-30111 Sequence 30111, A
c	95	180.5	19.7	565	14	US-10-198-846-366 Sequence 366, App
c	96	174	19.0	510	18	US-10-363-345A-653 Sequence 653, App
	97	174	19.0	510	18	US-10-363-345A-654 Sequence 654, App
c	98	174	19.0	510	19	US-10-363-483A-653 Sequence 653, App
	99	174	19.0	510	19	US-10-363-483A-654 Sequence 654, App
	100	172.5	18.8	1495	16	US-10-006-285-269 Sequence 269, App
	101	171.5	18.7	1627	14	US-10-116-802-322 Sequence 322, App
	102	171.5	18.7	2010	17	US-10-172-118-1482 Sequence 1482, Ap
	103	171.5	18.7	2010	17	US-10-342-887-1482 Sequence 1482, Ap
	104	163.5	17.8	2036	16	US-10-006-285-447 Sequence 447, App
	105	159.5	17.4	349	17	US-10-062-674-254 Sequence 254, App
	106	159	17.4	395	10	US-09-918-995-4376 Sequence 4376, Ap
	107	158	17.2	810	18	US-10-723-860-4925 Sequence 4925, Ap
	108	149.5	16.3	1891	16	US-10-316-253-108 Sequence 108, App
	109	148	16.2	403	9	US-09-960-352-3948 Sequence 3948, Ap
	110	148	16.2	490	17	US-10-369-493-30109 Sequence 30109, A
	111	148	16.2	490	17	US-10-369-493-30112 Sequence 30112, A
c	112	148	16.2	6806	17	US-10-205-331-105 Sequence 105, App
	113	146	15.9	442	10	US-09-918-995-5137 Sequence 5137, Ap
	114	142.5	15.6	510	18	US-10-363-345A-655 Sequence 655, App
c	115	142.5	15.6	510	18	US-10-363-345A-656 Sequence 656, App
	116	142.5	15.6	510	19	US-10-363-483A-655 Sequence 655, App
c	117	142.5	15.6	510	19	US-10-363-483A-656 Sequence 656, App
c	118	142	15.5	131	16	US-10-029-386-26899 Sequence 26899, A
c	119	142	15.5	451	15	US-10-102-524-483 Sequence 483, App
c	120	142	15.5	534	16	US-10-029-386-13199 Sequence 13199, A
	121	139	15.2	427	9	US-09-960-352-11415 Sequence 11415, A
	122	137.5	15.0	450	9	US-09-880-107-3022 Sequence 3022, Ap
	123	137.5	15.0	450	9	US-09-954-531-962 Sequence 962, App
	124	137.5	15.0	450	9	US-09-954-531-1355 Sequence 1355, Ap
	125	137.5	15.0	450	19	US-10-843-641A-2029 Sequence 2029, Ap
	126	137.5	15.0	450	19	US-10-843-641A-2422 Sequence 2422, Ap
	127	131	14.3	303	9	US-09-960-352-2010 Sequence 2010, Ap
	128	128	14.0	446	9	US-09-960-352-5348 Sequence 5348, Ap
	129	127.5	13.9	2153	17	US-10-062-674-2106 Sequence 2106, Ap
	130	125.5	13.7	495	14	US-10-255-536-69 Sequence 69, Appl
	131	124	13.5	423	9	US-09-833-790-52 Sequence 52, Appl
	132	123	13.4	339	9	US-09-867-701-6118 Sequence 6118, Ap
	133	123	13.4	468	9	US-09-938-842A-2558 Sequence 2558, Ap
	134	123	13.4	468	11	US-09-938-842A-2558 Sequence 2558, Ap
	135	121	13.2	65	10	US-09-908-975-30348 Sequence 30348, A
c	136	120.5	13.2	42339	13	US-10-087-192-991 Sequence 991, App
	137	115.5	12.6	478	10	US-09-918-995-20604 Sequence 20604, A
	138	115	12.6	1922	17	US-10-104-047-1686 Sequence 1686, Ap
c	139	115	12.6	50002	13	US-10-087-192-994 Sequence 994, App
c	140	114	12.4	932	18	US-10-425-115-183047 Sequence 183047,
	141	110.5	12.1	719	15	US-10-037-270-669 Sequence 669, App
	142	110.5	12.1	719	17	US-10-117-722-669 Sequence 669, App
	143	110.5	12.1	2152	18	US-10-723-860-395 Sequence 395, App
	144	110.5	12.1	2410	17	US-10-104-047-972 Sequence 972, App
	145	110.5	12.1	3503	18	US-10-723-860-5087 Sequence 5087, Ap
	146	110	12.0	599	18	US-10-021-323-5328 Sequence 5328, Ap
	147	107.5	11.7	462	9	US-09-938-842A-1962 Sequence 1962, Ap
	148	107.5	11.7	462	11	US-09-938-842A-1962 Sequence 1962, Ap
	149	106.5	11.6	817	18	US-10-767-701-13295 Sequence 13295, A
	150	106	11.6	3188	18	US-10-425-115-120839 Sequence 120839,

ALIGNMENTS

RESULT 1
US-10-316-253-112
; Sequence 112, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 112
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(534)
; OTHER INFORMATION:
US-10-316-253-112

Alignment Scores:

Pred. No.:	1.42e-108	Length:	1056
Score:	871.00	Matches:	164
Percent Similarity:	97.69%	Conservative:	5
Best Local Similarity:	94.80%	Mismatches:	4
Query Match:	95.09%	Indels:	0
DB:	16	Gaps:	0

US-10-657-740-1 (1-173) x US-10-316-253-112 (1-1056)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         13 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 72

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 132

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 192

Qy         61 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATCTCTGAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTGAAGCACTTC 252

Qy         81 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 100
              |||||||||||||||||||||||||||   :::||||||||||||||||||||||||||||
Db        253 TCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCATGGCAAACAC 312

Qy        101 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGCTACCGTCTG 372
```

```
Qy        121 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGCATGCTGACC 432

Qy        141 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 160
              |||   |||||||||:::|||:::||||||||||||   |||:::||||||||||||||||
Db        433 TTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGGGCCATTCCC 492

Qy        161 ValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||:::|||||||||||||||
Db        493 GTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 531

RESULT 2
US-10-316-253-110
; Sequence 110, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Peters, Kevin
;  APPLICANT:  Thompson, Larry
;  APPLICANT:  Wang, Feng
;  APPLICANT:  Greis, Kenneth
;  TITLE OF INVENTION: Angiogenesis Modulating Proteins
;  FILE REFERENCE: 8865M
;  CURRENT APPLICATION NUMBER: US/10/316,253
;  CURRENT FILING DATE:  2002-12-10
;  PRIOR APPLICATION NUMBER: US 60/355,295
;  PRIOR FILING DATE: 2002-02-08
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 110
;   LENGTH: 1271
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (159)..(749)
;   OTHER INFORMATION:
US-10-316-253-110

Alignment Scores:
Pred. No.:              1.65e-105       Length:         1271
Score:                  849.50          Matches:        164
Percent Similarity:     86.22%          Conservative:   5
Best Local Similarity:  83.67%          Mismatches:     4
Query Match:            92.74%          Indels:         23
DB:                     16              Gaps:           1

US-10-657-740-1 (1-173) x US-10-316-253-110 (1-1271)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        159 ATGGACGTCACCATCCAGCACCCTTGGTTCAAGCGCGCCCTGGGGCCCTTCTACCCCAGC 218

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGACTGTTCGACCAGTTCTTCGGCGAGGGCCTTTTTGAATACGACCTGCTGCCCTTCCTG 278

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCTTCCACCATCAGCCCCTACTACCGCCAGTCTCTCTTCCGCACAGTGTTGGACTCCGGC 338

Qy         61 IleSerGlu------------------------------------------------- 63
              |||||||||
Db        339 ATCTCTGAGCTCATGACCCATATGTGGTTTGTAATGCACCAACCACATGCTGGAAACCCC 398

Qy         64 ------------------ValArgSerAspArgAspLysPheValIlePheLeuAspVal 77
                                 ||||||||||||||||||||||||||||||||||||||||
Db        399 AAGAACAACCCCGGCAAGGTCCGATCTGACCGGGACAAGTTTGTCATCTTCTTGGATGTG 458

Qy         78 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 97
              ||||||||||||||||||||||||||||||||||   :::||||||||||||||||||||

Db        459 AAGCACTTCTCTCCTGAGGACCTCACCGTGAAGGTACTGGAAGATTTCGTGGAGATCCAT 518

Qy         98 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 GGCAAACACAACGAGAGGCAGGATGACCATGGCTACATTTCCCGTGAATTTCACCGTCGC 578

Qy        118 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 TACCGTCTGCCTTCCAATGTGGACCAGTCCGCCCTCTCCTGCTCCTTGTCTGCGGATGGC 638

Qy        138 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 157
              ||||||||||||   |||||||||:::|||:::||||||||||||   |||:::||||||
Db        639 ATGCTGACCTTCTCTGGCCCCAAGGTCCAGTCTGGCTTGGATGCTGGCCACAGCGAGAGG 698

Qy        158 AlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        699 GCCATTCCCGTGTCACGGGAGGAGAAGCCCAGCTCGGCACCCTCGTCC 746

RESULT 3
US-10-105-427-1
; Sequence 1, Application US/10105427
; Publication No. US20020177192A1
; GENERAL INFORMATION:
;  APPLICANT: Council of Scientific and Industrial Research
;  TITLE OF INVENTION: Chimeric protein alpha BNAC crystallin with extraordinarily high
;  TITLE OF INVENTION:  chaperone-like activity and a method thereof
;  FILE REFERENCE: US- 649
;  CURRENT APPLICATION NUMBER: US/10/105,427
;  CURRENT FILING DATE:  2002-07-03
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 531
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence for chimeric alpha BNAC
US-10-105-427-1

Alignment Scores:
Pred. No.:              7.01e-86        Length:         531
Score:                  702.00          Matches:        137
Percent Similarity:     85.96%          Conservative:   16
Best Local Similarity:  76.97%          Mismatches:     19
Query Match:            76.64%          Indels:         6
DB:                     13              Gaps:           4

US-10-657-740-1 (1-173) x US-10-105-427-1 (1-531)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db          1 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::         |||       |||      :::
Db        118 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        178 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 GATGTGAAGCACTTCTCCCCGGAGGACCTCACCGTGAAGGTGCAGGACGACTTTGTGGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ATCCACGGAAAGCACAACGAGCGCCAGGACGACCACGGCTACATTTCCCGTGAGTTCCAC 357
```

```
Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 CGCCGCTACCGCCTGCCGTCCAACGTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCC 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GATGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTGGATGCCACCCACGCC 477

Qy        156 GluArgAlaIleProValSerArgGluGluLysProThrSerAlaProSerSer 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 GAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAGCCCACCTCGGCTCCCTCGTCC 531

RESULT 4
US-10-657-740-2
; Sequence 2, Application US/10657740
; Publication No. US20040157289A1
; GENERAL INFORMATION:
;  APPLICANT: Salerno, John C.
;  APPLICANT:  Hanna, Michael
;  APPLICANT:  Koretz, Jane F.
;  APPLICANT:  Crone, Donna
;  APPLICANT:  Smith, Susan E.
;  TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM
;  FILE REFERENCE: 01794100H406US1
;  CURRENT APPLICATION NUMBER: US/10/657,740
;  CURRENT FILING DATE:  2003-09-08
;  PRIOR APPLICATION NUMBER: US 60/408,680
;  PRIOR FILING DATE: 2002-09-06
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-657-740-2

Alignment Scores:
Pred. No.:              1.74e-77      Length:         372
Score:                  639.00        Matches:        123
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            69.76%        Indels:         0
DB:                     18            Gaps:           0

US-10-657-740-1 (1-173) x US-10-657-740-2 (1-372)

Qy         51 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 70
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCCCTCTTCCGCACCGTGCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGGACAAG 60

Qy         71 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCCGGAGGACCTCACCGTGAAGGTGCAG 120

Qy         91 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACGACTTTGTGGAGATCCACGGAAAGCACAACGAGCGCCAGGACGACCACGGCTACATT 180

Qy        111 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAACGTGGACCAGTCGGCCCTCTCT 240

Qy        131 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTG 300

Qy        151 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGGGAGGAGAAGCCCACCTCGGCT 360

Qy        171 ProSerSer 173
              |||||||||
Db        361 CCCTCGTCC 369
```

```
RESULT 5
US-10-101-510-250
; Sequence 250, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
;  APPLICANT: WAN, JACKSON
;  APPLICANT:  WANG, YIXIN
;  TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
;  FILE REFERENCE: 15117.0012
;  CURRENT APPLICATION NUMBER: US/10/101,510
;  CURRENT FILING DATE:  2002-03-20
;  PRIOR APPLICATION NUMBER: 60/276,947
;  PRIOR FILING DATE: 2001-03-20
;  NUMBER OF SEQ ID NOS: 805
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 250
;   LENGTH: 548
;   TYPE: DNA
;   ORGANISM: Oryctolagus cuniculus
US-10-101-510-250

Alignment Scores:
Pred. No.:              3.06e-57      Length:         548
Score:                  493.00        Matches:        98
Percent Similarity:     74.16%        Conservative:   34
Best Local Similarity:  55.06%        Mismatches:     34
Query Match:            53.82%        Indels:         12
DB:                     15            Gaps:           6

US-10-657-740-1 (1-173) x US-10-101-510-250 (1-548)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         21 ATGGACATCGCTATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTTCACTCGCCC 80

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   |||||||   |||
Db         81 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTCTTCCCA--- 137

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        138 ACTTCTACTTCCCTGAGCCCCTTCTATCTTCGGCCACCCTCATTCCTGCGGGCACCCAGC 197

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        198 TGGATTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 257

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        258 GATGTGAAGCACTTCTCCCCAGAGGAGCTCAAGGTCAAAGTGTTGGGTGATGTGATTGAG 317

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||||
Db        318 GTGCACGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 377

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        378 AGGAAATACCGGATCCCAGCTGATGTGGACCCTCTCACCATTACTTCATCCCTGTCATCT 437

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||             |||
Db        438 GATGGGGTCCTCACTGTGAATGGACCAAGGAAGCAA-----------GCCCCTGGCCCA 485

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||       |||:::||||||
Db        486 GAGCGCACCATTCCCATAACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCC 539

RESULT 6
US-10-770-668-35
; Sequence 35, Application US/10770668
; Publication No. US20040191843A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wright, Susan C.
;  APPLICANT:  Larrick, James W.
;  APPLICANT:  Nock, Steffen R.
;  APPLICANT:  Wilson, David S.
;  TITLE OF INVENTION: Cell-Killing Molecules and Methods of Use Thereof
;  FILE REFERENCE: ABSALUS-08602
;  CURRENT APPLICATION NUMBER: US/10/770,668
;  CURRENT FILING DATE:  2004-02-02
;  NUMBER OF SEQ ID NOS: 81
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
;   LENGTH: 528
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-770-668-35

Alignment Scores:
Pred. No.:                1.02e-56      Length:          528
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       18            Gaps:            6

US-10-657-740-1 (1-173) x US-10-770-668-35 (1-528)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   ||||||||   :::|||         ||||||:::   |||
Db          1 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 60

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 117

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        118 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 177

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        178 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 237

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        238 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 297

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        298 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 357

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        358 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 417

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||||      ||||||:::   |||                 :::
Db        418 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 465

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 519

RESULT 7
US-09-954-456-514
; Sequence 514, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 514
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-514

Alignment Scores:
Pred. No.:                1.53e-56      Length:          691
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       9             Gaps:            6

US-10-657-740-1 (1-173) x US-09-954-456-514 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   ||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||||      ||||||:::   |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544
```

```
RESULT 8
US-09-960-706-869
; Sequence 869, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 869
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 S45630
US-09-960-706-869

Alignment Scores:
Pred. No.:                1.53e-56      Length:          691
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       10            Gaps:            6

US-10-657-740-1 (1-173) x US-09-960-706-869 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  |||||||||   :::|||         |||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||      |||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 9
```

```
US-09-873-319-566
; Sequence 566, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 566
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 S45630
US-09-873-319-566

Alignment Scores:
Pred. No.:                1.53e-56      Length:          691
Score:                    489.00        Matches:         97
Percent Similarity:       73.60%        Conservative:    34
Best Local Similarity:    54.49%        Mismatches:      35
Query Match:              53.38%        Indels:          12
DB:                       10            Gaps:            6

US-10-657-740-1 (1-173) x US-09-873-319-566 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  |||||||||   :::|||         |||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||      |||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 10
```

```
US-09-873-367C-764
; Sequence 764, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 764
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-764

Alignment Scores:
Pred. No.:              1.53e-56      Length:          691
Score:                  489.00        Matches:         97
Percent Similarity:     73.60%        Conservative:    34
Best Local Similarity:  54.49%        Mismatches:      35
Query Match:            53.38%        Indels:          12
DB:                     10            Gaps:            6

US-10-657-740-1 (1-173) x US-09-873-367C-764 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||  |||  |||  ||||||  |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||  |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||||      |||||||:::  |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 11
US-10-133-937-61
; Sequence 61, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
;  APPLICANT: Khan, Javed
;  APPLICANT:  Ringner, Markus
;  APPLICANT:  Peterson, Carsten
;  APPLICANT:  Meltzer, Paul
;  TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
;  TITLE OF INVENTION:  DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
;  TITLE OF INVENTION:  OTHER BIOLOGICAL STATES
;  FILE REFERENCE: 11613.56US01
;  CURRENT APPLICATION NUMBER: US/10/133,937
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 99
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-133-937-61

Alignment Scores:
Pred. No.:              1.53e-56      Length:          691
Score:                  489.00        Matches:         97
Percent Similarity:     73.60%        Conservative:    34
Best Local Similarity:  54.49%        Mismatches:      35
Query Match:            53.38%        Indels:          12
DB:                     17            Gaps:            6

US-10-657-740-1 (1-173) x US-10-133-937-61 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||  |||  |||  ||||||  |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||  |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::|||||||      |||||||:::  |||                 :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544
```

```
RESULT 12
US-10-172-118-686
; Sequence 686, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 686
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: NM_001885
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-686

Alignment Scores:
Pred. No.:              1.53e-56      Length:         691
Score:                  489.00        Matches:        97
Percent Similarity:     73.60%        Conservative:   34
Best Local Similarity:  54.49%        Mismatches:     35
Query Match:            53.38%        Indels:         12
DB:                     17            Gaps:           6

US-10-657-740-1 (1-173) x US-10-172-118-686 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              |||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::||||||         ::::::   ||||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||

Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 13
US-10-159-563-61
; Sequence 61, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
;  APPLICANT: Khan, Javed
;  APPLICANT:  Ringner, Markus
;  APPLICANT:  Peterson, Carsten
;  APPLICANT:  Meltzer, Paul
;  TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
;  TITLE OF INVENTION:  DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
;  FILE REFERENCE: 11613.56USI1
;  CURRENT APPLICATION NUMBER: US/10/159,563
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US 10/133,937
;  PRIOR FILING DATE: 2002-04-25
;  NUMBER OF SEQ ID NOS: 444
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-159-563-61

Alignment Scores:
Pred. No.:              1.53e-56      Length:         691
Score:                  489.00        Matches:        97
Percent Similarity:     73.60%        Conservative:   34
Best Local Similarity:  54.49%        Mismatches:     35
Query Match:            53.38%        Indels:         12
DB:                     17            Gaps:           6

US-10-657-740-1 (1-173) x US-10-159-563-61 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||:::::::::|||||:::|||   :::       |||       |||       :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              |||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::||||||         ::::::   ||||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG------------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 14
```

```
US-10-342-887-686
; Sequence 686, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter S.
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Christopher J.
;  APPLICANT:  Van 't Veer, Laura Johanna
;  APPLICANT:  Van de Vijver, Marc J.
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 686
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-686

Alignment Scores:
Pred. No.:              1.53e-56      Length:          691
Score:                  489.00        Matches:         97
Percent Similarity:     73.60%        Conservative:    34
Best Local Similarity:  54.49%        Mismatches:      35
Query Match:            53.38%        Indels:          12
DB:                     17            Gaps:            6

US-10-657-740-1 (1-173) x US-10-342-887-686 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||::::::::::||||||:::|||   :::      |||      |||   :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::::||||||          ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 15
US-10-843-641A-764
; Sequence 764, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 764
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-764

Alignment Scores:
Pred. No.:              1.53e-56      Length:          691
Score:                  489.00        Matches:         97
Percent Similarity:     73.60%        Conservative:    34
Best Local Similarity:  54.49%        Mismatches:      35
Query Match:            53.38%        Indels:          12
DB:                     19            Gaps:            6

US-10-657-740-1 (1-173) x US-10-843-641A-764 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 |||::::::::::||||||:::|||   :::      |||      |||   :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                 |||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
```

```
                :::||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   ||||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 16
US-10-843-641A-3541
; Sequence 3541, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3541
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-3541

Alignment Scores:
Pred. No.:              1.53e-56        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     19              Gaps:           6

US-10-657-740-1 (1-173) x US-10-843-641A-3541 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              |||:::|||:::||||||::|||   ::::::|||:::|||   :::   |||
Db        203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              |||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::||||||:::||||||||||||||||||
Db        323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||::::::||||||         ::::::   |||||||||:::
Db        383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||       ||||||:::   |||                  :::
Db        443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544

RESULT 17
US-10-367-057-187
; Sequence 187, Application US/10367057
; Publication No. US20050100554A1
; GENERAL INFORMATION:
;  APPLICANT: Cuthill, Scott;
;  APPLICANT:  Jackson, Amanda;
;  APPLICANT:  Lewin, David A.;
;  APPLICANT:  Ooi, Chean Eng
;  TITLE OF INVENTION: Complexes and Methods of Using Same
;  FILE REFERENCE: 21402-559
;  CURRENT APPLICATION NUMBER: US/10/367,057
;  CURRENT FILING DATE:  2003-02-14
;  PRIOR APPLICATION NUMBER: 60/256,911
;  PRIOR FILING DATE: 2002-02-14
;  NUMBER OF SEQ ID NOS: 198
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 187
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-367-057-187

Alignment Scores:
Pred. No.:              1.53e-56        Length:         691
Score:                  489.00          Matches:        97
Percent Similarity:     73.60%          Conservative:   34
Best Local Similarity:  54.49%          Mismatches:     35
Query Match:            53.38%          Indels:         12
DB:                     19              Gaps:           6

US-10-657-740-1 (1-173) x US-10-367-057-187 (1-691)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         26 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 85

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         86 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 142

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              |||:::::::::||||||:::|||   :::      |||      |||      :::
Db        143 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 202
```

```
Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db         203 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 262

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db         263 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 322

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db         323 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 382

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::|||||         ::::::   |||||||||:::
Db         383 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTCACCATTACTTCATCCCTGTCATCT 442

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::|||||||      ||||||:::   |||                 :::
Db         443 GATGGGGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GTCTCTGGCCCT 490

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db         491 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACCGCAGCCCCC 544
```

RESULT 18
US-10-152-319A-1574
; Sequence 1574, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1574
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_012935
US-10-152-319A-1574

Alignment Scores:

Pred. No.:	3.6e-56	Length:	528
Score:	485.00	Matches:	96
Percent Similarity:	74.16%	Conservative:	36
Best Local Similarity:	53.93%	Mismatches:	34
Query Match:	52.95%	Indels:	12
DB:	17	Gaps:	6

US-10-657-740-1 (1-173) x US-10-152-319A-1574 (1-528)

```
Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||  |||||||||   :::|||         ||||||:::   |||
Db           1 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 60

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               |||||||||||||||||||||||||||||||   |||   |||   ||||||
Db          61 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 117

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
               :::::::::::|||||:::|||  :::      |||      |||      :::
Db         118 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 177

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         178 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 237

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db         238 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 297

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db         298 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 357

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::::|||||         ::::::   |||||||||:::
Db         358 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 417

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||      |||||||:::   |||                 |||:::
Db         418 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 465

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||:::::|||||||||||||||      |||:::||||||
Db         466 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 519
```

RESULT 19
US-10-486-706-455
; Sequence 455, Application US/10486706
; Publication No. US20050071088A1
; GENERAL INFORMATION:
; APPLICANT: LANDFIELD, PHILIP W.
; APPLICANT: BLALOCK, ERIC M.
; APPLICANT: CHEN, KUEY-CHU
; APPLICANT: FOSTER, THOMAS C.
; TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR
; TITLE OF INVENTION: BRAIN AGING AND AGE-RELATED COGNITIVE IMPAIRMENT
; FILE REFERENCE: 50229-426
; CURRENT APPLICATION NUMBER: US/10/486,706
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: PCT/US02/25607
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/311,343
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 455
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Rattus rattus
US-10-486-706-455

Alignment Scores:

Pred. No.:	5.37e-56	Length:	689
Score:	485.00	Matches:	96
Percent Similarity:	74.16%	Conservative:	36

```
Best Local Similarity:  53.93%          Mismatches:    34
Query Match:            52.95%          Indels:        12
DB:                     19              Gaps:          6

US-10-657-740-1 (1-173) x US-10-486-706-455 (1-689)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  ||||||||   :::|||         ||||||:::   |||
Db         13 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 72

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||
Db         73 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTTTCT--- 129

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 ::::::::::::|||||:::|||   :::      |||      |||      :::
Db        130 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 189

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        190 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 249

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        250 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 309

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        310 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 369

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::||||||         ::::::   ||||||||||:::
Db        370 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 429

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||            |||:::
Db        430 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 477

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db        478 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 531

RESULT 20
US-09-917-800A-1419
; Sequence 1419, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1419
;   LENGTH: 1247
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 M55534
US-09-917-800A-1419

Alignment Scores:
Pred. No.:              1.31e-55        Length:        1247
Score:                  485.00          Matches:       96
Percent Similarity:     74.16%          Conservative:  36
Best Local Similarity:  53.93%          Mismatches:    34
Query Match:            52.95%          Indels:        12
DB:                     9               Gaps:          6

US-10-657-740-1 (1-173) x US-09-917-800A-1419 (1-1247)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||  ||||||||   :::|||         ||||||:::   |||
Db        583 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 642

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||
Db        643 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 699

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 ::::::::::::|||||:::|||   :::      |||      |||      :::
Db        700 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 759

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db        760 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 819

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        820 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 879

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||
Db        880 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 939

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
              |||:::||||||:::|||:::||||||         ::::::   ||||||||||:::
Db        940 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 999

Qy        136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
              ||||||:::||||||      ||||||:::   |||            |||:::
Db       1000 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 1047

Qy        156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
              ||||||   ||||||::::::|||||||||||||||      |||:::||||||
Db       1048 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 1101

RESULT 21
US-10-486-706-105
; Sequence 105, Application US/10486706
; Publication No. US20050071088A1
; GENERAL INFORMATION:
;  APPLICANT: LANDFIELD, PHILIP W.
;  APPLICANT:  BLALOCK, ERIC M.
;  APPLICANT:  CHEN, KUEY-CHU
;  APPLICANT:  FOSTER, THOMAS C.
;  TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR
;  TITLE OF INVENTION:  BRAIN AGING AND AGE-RELATED COGNITIVE IMPAIRMENT
;  FILE REFERENCE: 50229-426
;  CURRENT APPLICATION NUMBER: US/10/486,706
;  CURRENT FILING DATE:  2004-02-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25607
```

```
;  PRIOR FILING DATE: 2002-08-13
;  PRIOR APPLICATION NUMBER: US 60/311,343
;  PRIOR FILING DATE: 2001-08-13
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 105
;   LENGTH: 1247
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-10-486-706-105

Alignment Scores:
Pred. No.:              1.31e-55        Length:         1247
Score:                  485.00          Matches:        96
Percent Similarity:     74.16%          Conservative:   36
Best Local Similarity:  53.93%          Mismatches:     34
Query Match:            52.95%          Indels:         12
DB:                     19              Gaps:           6

US-10-657-740-1 (1-173) x US-10-486-706-105 (1-1247)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         583 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 642

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               ||||||||||||||||||||||||||||||   |||   |||   ||||||
Db         643 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 699

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  ::::::::::::||||||:::|||   :::       |||       |||       :::
Db         700 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 759

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                  :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         760 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 819

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db         820 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 879

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||:::
Db         880 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 939

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::||||||         ::::::   ||||||||||:::
Db         940 AGGAAGTACCGGATCCCAGCCGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCATCG 999

Qy         136 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 155
               ||||||:::||||||       ||||||:::   |||             |||:::
Db        1000 GATGGAGTCCTCACTGTGAATGGACCAAGGAAACAG-----------GCCTCTGGCCCT 1047

Qy         156 GluArgAlaIleProValSerArgGluGluLysPro------ThrSerAlaPro 171
               ||||||   ||||||::::::||||||||||||||||      |||:::||||||
Db        1048 GAGCGCACCATTCCCATCACCCGTGAAGAGAAGCCTGCTGTCACTGCAGCCCCT 1101

RESULT 22
US-09-960-352-4277
; Sequence 4277, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4277
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 19-BOVMS1-021-Q1-E1-E3
US-09-960-352-4277

Alignment Scores:
Pred. No.:              4.22e-47        Length:         449
Score:                  418.00          Matches:        79
Percent Similarity:     77.30%          Conservative:   30
Best Local Similarity:  56.03%          Mismatches:     26
Query Match:            45.63%          Indels:         6
DB:                     9               Gaps:           4

US-10-657-740-1 (1-173) x US-09-960-352-4277 (1-449)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          30 ATGGATATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTCCCTTTCCACTCTCCC 89

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               ||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db          90 AGCCGCCTCTTTGACCAGTTTTTTGGCGAGCACCTGTTAGAGTCTGATCTCTTCCCA--- 146

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  |||:::::::::||||||:::|||   :::       |||       |||       :::
Db         147 GCTTCTACTTCCCTGAGCCCCTTCTACCTTCGGCCGCCCTCATTTCTGCGGGCACCCAGC 206

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                  :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db         207 TGGATTGACACTGGCCTCTCAGAGATGCGTCTGGAGAAGGACAGATTCTCTGTCAACCTG 266

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db         267 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAGGTCAAGGTGCTGGGAGATGTGATTGAG 326

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
               :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||:::
Db         327 GTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGGGAGTTCCAC 386

Qy         116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 135
               |||:::||||||:::|||:::::||||||         ::::::   ||||||||||:::
Db         387 AGGAAATACCGGATCCCAGCTGACGTGGAACCTCTCGCCATTACTTCATCCCTGTCGTCT 446

Qy         136 Asp 136
               |||
Db         447 GAT 449

RESULT 23
US-09-960-352-12622
; Sequence 12622, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12622
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 54-LIB3057-007-Q1-K1-F6
US-09-960-352-12622

Alignment Scores:
Pred. No.:              6.75e-44        Length:         396
Score:                  394.00          Matches:        74
```

```
Percent Similarity:     77.52%          Conservative:  26
Best Local Similarity:  57.36%          Mismatches:    23
Query Match:            43.01%          Indels:        6
DB:                     9               Gaps:          4

US-10-657-740-1 (1-173) x US-09-960-352-12622 (1-396)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         13 ATGGATATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTCCCTTTCCACTCTCCC 72

Qy         20 SerArgLeuPheAspGlnPheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         73 AGCCGCCTCTTTGACCAGTTTTTTGGCGAGCACCTGTTGGAGTCTGATCTCTTCCCA--- 129

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        130 GCTTCTACTTCCCTGAGCCCCTTCTACCTTCGGCCGCCCTCATTTCTGCGGGCACCCAGC 189

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        190 TGGATTGACACTGGCCTCTCAGAGATGCGTCTGGAGAAGGACAGATTCTCTGTCAACCTG 249

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db        250 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAGGTCAAGGTGCTGGGAGATGTGATTGAG 309

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||   |||||||||||:::|||||:::|||||||||||||||||||
Db        310 GTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGGGAGTTCCAC 369

Qy        116 ArgArgTyrArgLeuProSerAsnVal 124
              |||:::||||||:::|||::::::|||
Db        370 AGGAAATACCGGATCCCAGCTGACGTG 396

RESULT 24
US-10-029-386-26581
; Sequence 26581, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26581
;   LENGTH: 211
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR21.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
;   OTHER INFORMATION: SWISSPROT HIT: P02489, EVALUE 3.00e-36
;   OTHER INFORMATION: EST_HUMAN HIT: BF726399.1, EVALUE 1.00e-113
;   OTHER INFORMATION: NT HIT: gi14780619, EVALUE 1.00e-115
US-10-029-386-26581

Alignment Scores:
Pred. No.:              1.31e-40        Length:        211
Score:                  367.00          Matches:       70
Percent Similarity:     100.00%         Conservative:  0
Best Local Similarity:  100.00%         Mismatches:    0
Query Match:            40.07%          Indels:        0
DB:                     16              Gaps:          0

US-10-657-740-1 (1-173) x US-10-029-386-26581 (1-211)

Qy         104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAAC 60

Qy         124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGC 120

Qy         144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGG 180

Qy         164 GluGluLysProThrSerAlaProSerSer 173
               ||||||||||||||||||||||||||||
Db         181 GAGGAGAAGCCCACCTCGGCTCCCTCGTCC 210

RESULT 25
US-10-029-386-12881
; Sequence 12881, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12881
;   LENGTH: 573
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR21.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
;   OTHER INFORMATION: NT HIT: AF026952.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF726856.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P02489, EVALUE 1.00e-32
US-10-029-386-12881

Alignment Scores:
Pred. No.:              5.86e-40        Length:        573
Score:                  367.00          Matches:       70
Percent Similarity:     100.00%         Conservative:  0
Best Local Similarity:  100.00%         Mismatches:    0
Query Match:            40.07%          Indels:        0
DB:                     16              Gaps:          0

US-10-657-740-1 (1-173) x US-10-029-386-12881 (1-573)

Qy         104 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 123
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         203 CAGGACGACCACGGCTACATTTCCCGTGAGTTCCACCGCCGCTACCGCCTGCCGTCCAAC 262

Qy         124 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         263 GTGGACCAGTCGGCCCTCTCTTGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGC 322

Qy         144 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         323 CCCAAGATCCAGACTGGCCTGGATGCCACCCACGCCGAGCGAGCCATCCCCGTGTCGCGG 382

Qy         164 GluGluLysProThrSerAlaProSerSer 173
               ||||||||||||||||||||||||||||
Db         383 GAGGAGAAGCCCACCTCGGCTCCCTCGTCC 412
```

```
RESULT 26
US-10-198-846-10360
; Sequence 10360, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10360
;   LENGTH: 411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-846-10360

Alignment Scores:
Pred. No.:                7.82e-40        Length:          411
Score:                    364.50          Matches:         71
Percent Similarity:       75.40%          Conservative:    24
Best Local Similarity:    56.35%          Mismatches:      24
Query Match:              39.79%          Indels:          7
DB:                       14              Gaps:            5

US-10-657-740-1 (1-173) x US-10-198-846-10360 (1-411)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db         39 ATGGACATAGCCATCCACCACCCCTGGATCCGGCGTCCCTTCTTTCCTTTCCACTCCCCA 98

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||||||||||||||||   |||   |||   ||||||
Db         99 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGACCTCTTCTCT--- 155

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                 ::::::::::|||||:::|||   :::      |||      |||      :::
Db        156 ACAGCCACTTCCCTGAGCCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCTAGC 215

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::||||||:::|||   ::::::|||:::|||   :::   |||
Db        216 TGGATTGACACTGGGCTCTCAGAGATGCGTATGGAGAAGGACAGGTTCTCTGTGAACCTG 275

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||||||||       |||   :::|||
Db        276 GACGTGAAGCACTTCTCTCCAGAGGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAG 335

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||   ||||||||||:::||||||:::||||||||||||||||||
Db        336 GTGCACGGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGGAGTTCCAC 395

Qy        116 ArgArgTyrArgLeuPro 121
              |||:::|||   ||||||
Db        396 AGGAAGTAC---CTGCCC 410

RESULT 27
US-09-918-995-4949
; Sequence 4949, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4949
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(450)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-4949

Alignment Scores:
Pred. No.:                2.7e-39         Length:          450
Score:                    361.00          Matches:         73
Percent Similarity:       71.94%          Conservative:    27
Best Local Similarity:    52.52%          Mismatches:      33
Query Match:              39.41%          Indels:          6
DB:                       10              Gaps:            4

US-10-657-740-1 (1-173) x US-09-918-995-4949 (1-450)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
              ||||||:::   |||   |||||||   :::|||         ||||:::   |||
Db         37 ATGGACATCGCCATCCACCGCCCCTGGATCCGCCGACCCTTCTTTGCTTTCCACTCCCCC 96

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||||||||||||   |||||||||   |||   |||   ||||||   |||
Db         97 AGCCGCCTCTTTGACCAGTGCTTCGGAGAGCACCTGTCGGAGTCTGATCTTTGCCCG--- 153

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGlnSer------LeuPheArg------Thr 55
                 |||::::::::::||||||:::|||   :::|||       |||   |||       :::
Db        154 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGTCACCCTGCTTACTGCGGGCACCCAGC 213

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              |||:::|||:::|||::::::|||   ::::::|||:::      :::   |||
Db        214 TGGTGTGACACTGGACTCTCAAAGATGCGCCTGGAGAAGGACAGGGTCTCTGTCAACCTG 273

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::|||   |||   |||       |||   :::|||
Db        274 GATGTGAAGCACTTCTCCCCAGAGGAACTCAGAGTTAATGTGTTGTGAGATGAGATTGAG 333

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::|||||||||||||   ||||||||||||:::||||||:::||||||||||||||||||
Db        334 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 393

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
              |||   ||||||:::|||:::::::||||||         ::::::   |||||||||
Db        394 AGGACATACCGGATCCCAGCTGATGTAGACCCTCTGACCATTACTTCATCCCTGTCN 450

RESULT 28
US-10-029-386-1061
; Sequence 1061, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1061
;   LENGTH: 579
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR21_69.0
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.84
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
;   OTHER INFORMATION: SWISSPROT HIT: P02488, EVALUE 5.00e-33
;   OTHER INFORMATION: NT HIT: gi14780622, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF726222.1, EVALUE 0.00e+00
US-10-029-386-1061

Alignment Scores:
Pred. No.:              3.26e-37        Length:         579
Score:                  347.00          Matches:        65
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            37.88%          Indels:         0
DB:                     16              Gaps:           0

US-10-657-740-1 (1-173) x US-10-029-386-1061 (1-579)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 205

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 265

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 325

Qy         61 IleSerGluValArg 65
              |||||||||||||||
Db        326 ATCTCTGAGGTAAGA 340

RESULT 29
US-10-027-632-172414
; Sequence 172414, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 172414
;   LENGTH: 834
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-172414

Alignment Scores:
Pred. No.:              5.64e-37        Length:         834
Score:                  347.00          Matches:        65
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            37.88%          Indels:         0
DB:                     13              Gaps:           0

US-10-657-740-1 (1-173) x US-10-027-632-172414 (1-834)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATGGAYGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 310

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 370

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 430

Qy         61 IleSerGluValArg 65
              |||||||||||||||
Db        431 ATCTCTGAGGTAAGA 445

RESULT 30
US-10-027-632-172414
; Sequence 172414, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 172414
;   LENGTH: 834
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-172414

Alignment Scores:
Pred. No.:              5.64e-37        Length:         834
Score:                  347.00          Matches:        65
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            37.88%          Indels:         0
DB:                     17              Gaps:           0

US-10-657-740-1 (1-173) x US-10-027-632-172414 (1-834)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATGGAYGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 310
```

```
Qy          21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         311 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 370

Qy          41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         371 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 430

Qy          61 IleSerGluValArg 65
               |||||||||||||||
Db         431 ATCTCTGAGGTAAGA 445

RESULT 31
US-10-062-674-1924
; Sequence 1924, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
;  APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
;  TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
;  FILE REFERENCE: PA-0026-1 CIP
;  CURRENT APPLICATION NUMBER: US/10/062,674
;  CURRENT FILING DATE:  2002-01-30
;  PRIOR APPLICATION NUMBER: US 09/625,102
;  PRIOR FILING DATE: 2000-07-24
;  NUMBER OF SEQ ID NOS: 2217
;  SOFTWARE: PERL Program
; SEQ ID NO 1924
;   LENGTH: 829
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20040005559A1 347789.24
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1) ... (829)
;   OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-1924

Alignment Scores:
Pred. No.:              8.97e-37       Length:         829
Score:                  345.50         Matches:        76
Percent Similarity:     72.41%         Conservative:   29
Best Local Similarity:  52.41%         Mismatches:     30
Query Match:            37.72%         Indels:         10
DB:                     17             Gaps:           5

US-10-657-740-1 (1-173) x US-10-062-674-1924 (1-829)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||          ||||||:::   |||
Db         400 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 459

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               |||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db         460 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 516

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
               |||:::::::::||||||:::|||   :::      |||      |||      :::
Db         517 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 576

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
               |||:::|||:::||||||:::|||   :::::::|||:::|||   :::   |||
Db         577 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 636

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               ||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         637 GATGTGAAGCACTTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAG 696

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArg---GluPhe 114
               :::||||||||||   ||||||||||:::|||||:::|||||||||
Db         697 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGGGGNTNCN 756

Qy         115 HisArg-ArgTyrArg-LeuProSerAsnValAspGln-SerAlaLeuSerCysSerLeu 133
               ||| :::|||||| :::|||:::::|||||    |||   ::::::   ||||||
Db         757 ACCAGGGAAATACCGGGATCCCAGCTGATGTAGACCCTCTCAACCATTACTTCATCCCTG 816

Qy         134 SerAlaAsp 136
               |||::::::
Db         817 TCATCTGAA 825

RESULT 32
US-10-226-956-320
; Sequence 320, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
;  APPLICANT: Brophy, Colleen
;  APPLICANT:  Komalavilas, Padmini
;  APPLICANT:  Panitch, Alyssa
;  APPLICANT:  Joshi, Lokesh
;  APPLICANT:  Seal, Brandon L.
;  TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
;  FILE REFERENCE: ASU-1061-US
;  CURRENT APPLICATION NUMBER: US/10/226,956
;  CURRENT FILING DATE:  2002-08-23
;  PRIOR APPLICATION NUMBER: 60/314,535
;  PRIOR FILING DATE: 2001-08-23
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 320
;   LENGTH: 480
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Sequence encoding human HSP20
US-10-226-956-320

Alignment Scores:
Pred. No.:              2.24e-36       Length:         480
Score:                  340.00         Matches:        69
Percent Similarity:     58.14%         Conservative:   31
Best Local Similarity:  40.12%         Mismatches:     50
Query Match:            37.12%         Indels:         22
DB:                     14             Gaps:           3

US-10-657-740-1 (1-173) x US-10-226-956-320 (1-480)

Qy           3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
               |||   :::|||        |||   :::|||        |||
Db          13 GTTCCAGTCCAGCCTAGTTGGCTAAGAAGAGCTAGTGCGCCTTTGCCGGGTTTGAGTGCC 72

Qy          19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
               |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db          73 CCCGGGAGGCTATTTGATCAACGCTTTGGCGAGGGGTTACTGGAGGCTGAATTAGCAGCA 132

Qy          39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                  :::|||:::::|||||||||          |||              :::
Db         133 CTTTGTCCGACCACACTCGCGCCCTATTAC---------CTTAGAGCGCCGTCTGTAGCC 183

Qy          59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  :::::::::|||   :::|||          |||   :::   ||||||||||||
Db         184 TTACCAGTCGCTCAGGTACCAACTGACCCAGGCCACTTCTCCGTTTTATTAGACGTGAAA 243

Qy          79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
               |||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db         244 CACTTTAGCCCAGAAGAGATAGCAGTCAAAGTTGTAGGAGAGCATGTGGAAGTTCACGCG 303

Qy          99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
               :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||
Db         304 AGACATGAAGAGAGACCAGATGAACATGGTTTCGTAGCGAGAGAATTCCATCGACGGTAT 363

Qy         119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
               |||||||||      ||||||   :::|||::::::   :::||||||   :::|||:::
Db         364 CGTCTGCCCCCAGGAGTCGATCCTGCAGCTGTGACGAGTGCATTATCGCCTGAGGGAGTG 423

Qy         139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
```

```
Db        424 CTCAGTATC-----------------------------------------------CAAGCA 438

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170
              |||   |||   :::     |||   :::|||
Db        439 GCCCCCGCGTCAGCCCAAGCCCCGCCTCCGGCTGCT 474
```

RESULT 33
US-10-723-860-5075
; Sequence 5075, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5075
; LENGTH: 1820
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5075

Alignment Scores:

Pred. No.:	1.66e-35	Length:	1820
Score:	340.00	Matches:	69
Percent Similarity:	58.14%	Conservative:	31
Best Local Similarity:	40.12%	Mismatches:	50
Query Match:	37.12%	Indels:	22
DB:	18	Gaps:	3

US-10-657-740-1 (1-173) x US-10-723-860-5075 (1-1820)

```
Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
              |||   :::|||     |||   :::|||         |||
Db         35 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 94

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   |||||||||||||||   ||||||||||||||   |||   :::|||
Db         95 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 154

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                    :::|||::::::|||||||||         |||           :::
Db        155 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 205

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                  :::::::::|||   :::|||         |||   :::   ||||||||||||
Db        206 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 265

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   |||||:::|||
Db        266 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 325

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::|||||:::::::::|||||||||||||||||||
Db        326 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 385

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              |||||||||      ||||||   :::|||::::::   :::||||||   :::|||:::
Db        386 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACGTCCGCGCTGTCCCCCGAGGGCGTC 445

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                           :::|||
Db        446 CTGTCCATC-----------------------------------------CAGGCC 460

Qy        159 IleProValSerArgGluGluLysProThrSerAla 170

```
              |||   |||   :::     |||   :::|||
Db        461 GCACCAGCGTCGGCCCAGGCCCCACCGCCAGCCGCA 496
```

RESULT 34
US-10-029-386-14764
; Sequence 14764, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14764
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21_69.0
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: NT HIT: U05569.1, EVALUE 1.00e-116
; OTHER INFORMATION: EST_HUMAN HIT: BF727348.1, EVALUE 1.00e-116
; OTHER INFORMATION: SWISSPROT HIT: P02493, EVALUE 9.00e-33
US-10-029-386-14764

Alignment Scores:

Pred. No.:	1.23e-36	Length:	212
Score:	338.00	Matches:	63
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	36.90%	Indels:	0
DB:	16	Gaps:	0

US-10-657-740-1 (1-173) x US-10-029-386-14764 (1-212)

```
Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrProSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGGACGTGACCATCCAGCACCCCTGGTTCAAGCGCACCCTGGGGCCCTTCTACCCCAGC 83

Qy         21 ArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 CGGCTGTTCGACCAGTTTTTCGGCGAGGGCCTTTTTGAGTATGACCTGCTGCCCTTCCTG 143

Qy         41 SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TCGTCCACCATCAGCCCCTACTACCGCCAGTCCCTCTTCCGCACCGTGCTGGACTCCGGC 203

Qy         61 IleSerGlu 63
              ||||||||
Db        204 ATCTCTGAG 212
```

RESULT 35
US-09-918-995-4772
; Sequence 4772, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
```

```
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4772
;   LENGTH: 421
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-4772

Alignment Scores:
Pred. No.:                3.45e-36        Length:         421
Score:                    338.00          Matches:        68
Percent Similarity:       72.26%          Conservative:   31
Best Local Similarity:    49.64%          Mismatches:     32
Query Match:              36.90%          Indels:         7
DB:                       10              Gaps:           4

US-10-657-740-1 (1-173) x US-09-918-995-4772 (1-421)

Qy          1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyrPro--- 19
              ||||||:::   |||   |||||||||   :::|||          ||||||:::
Db         15 ATGGACATCGCCATCCACCACCCTTGGATCCGCCGGCCCTTCTTTCCTTTTCACTCCCGC 74

Qy         20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
              ||||||   |||   ||||||||||||   |||   |||   |||||   |||
Db         75 GGCCGCCTCTTAGACCATTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 131

Qy         40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
              |||:::::::::|||||:::|||   :::      |||      |||      :::
Db        132 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTTCTGCGGGCACCCAGC 191

Qy         56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
              :::|||:::|||:::|||:::::::|||   ::::::||||||      :::   |||
Db        192 TGGGTTGACACTGGACTCTCAAAGATGCGCCTGGAGAAGGACAAGATCTCTGTCAACCTG 251

Qy         76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
              ||||||||||||||||||||||:::::|||      ||||||      |||   ::::::
Db        252 GATGTGAAGCACTTCTCCCCAAAGGAACTCAAAGGTAAGGTGTG-GGAGACGTGATTGAC 310

Qy         96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHis 115
              :::||||||||||||:::|||||||||||:::||||||:::||||||||||||||||||
Db        311 GTGCATGGTAAACATGACGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCAC 370

Qy        116 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 132
              |||:::||||||:::|||::::::||||||      ::::::   |||
Db        371 AGGAAATACCGGATCCCAGCTGATGTAGACCCTCTTACCATTACTTCATCC 421

RESULT 36
US-10-226-956-303
; Sequence 303, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
;  APPLICANT: Brophy, Colleen
;  APPLICANT:  Komalavilas, Padmini
;  APPLICANT:  Panitch, Alyssa
;  APPLICANT:  Joshi, Lokesh
;  APPLICANT:  Seal, Brandon L.
;  TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
;  FILE REFERENCE: ASU-1061-US
;  CURRENT APPLICATION NUMBER: US/10/226,956
;  CURRENT FILING DATE:  2002-08-23
;  PRIOR APPLICATION NUMBER: 60/314,535
;  PRIOR FILING DATE: 2001-08-23
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 303
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Sequence encoding rat  HSP20
US-10-226-956-303

Alignment Scores:
Pred. No.:                8.04e-36        Length:         486
Score:                    336.00          Matches:        71
Percent Similarity:       57.14%          Conservative:   29
Best Local Similarity:    40.57%          Mismatches:     53
Query Match:              36.68%          Indels:         22
DB:                       14              Gaps:           3

US-10-657-740-1 (1-173) x US-10-226-956-303 (1-486)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe----------Tyr 18
              |||   :::|||      |||   :::|||          |||
Db         13 GTGCCGGTACAACCCAGCTGGCTGCGGCGTGCTTCCGCGCCATTACCTGGCTTCAGTACC 72

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db         73 CCCGGACGATTGTTTGACCAGAGGTTTGGGGAAGGTTTACTTGAGGCGGAATTAGCAAGT 132

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                     :::   |||:::|||||||||          |||              :::
Db        133 CTATGTCCTGCAGCTATAGCACCCTACTAC---------CTAAGGGCACCATCTGTCGCG 183

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                     ::::::|||   :::|||      |||   :::   ||||||||||
Db        184 CTCCCAACTGCCCAAGTGCCCACGGATCCAGGCTATTTCAGCGTTCTGTTAGACGTAAAG 243

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||:::::::::||||||||      |||   ||||||:::|||
Db        244 CATTTTAGTCCAGAAGAAATTTCAGTAAAAGTAGTGGGAGACCATGTCGAGGTACATGCT 303

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::|||:::|||||||||||||||||||
Db        304 AGACACGAAGAGAGACCTGATGAACACGGTTTCATCGCTCGAGAGTTTCACCGGCGTTAT 363

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
              ||||||||      ||||||   :::|||::::::   :::||||||   :::|||:::
Db        364 CGCTTGCCGCCGGGGGTTGATCCCGCGGCCGTCACATCAGCACTCAGTCCGGAGGGAGTT 423

Qy        139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
              |||:::                                                :::|||
Db        424 TTATCCATA----------------------------------------------CAAGCC 438

Qy        159 IleProValSerArgGluGluLysProThrSerAlaProSerSer 173
                 |||   |||   :::            |||   |||::::::
Db        439 ACACCGGCCTCTGCTCAGGCCTCGCTTCCATCGCCTCCTGCGGCA 483

RESULT 37
US-10-152-319A-1412
; Sequence 1412, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Higgs, Brandon
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5089-US
;  CURRENT APPLICATION NUMBER: US/10/152,319A
;  CURRENT FILING DATE:  2002-05-22
;  PRIOR APPLICATION NUMBER: US 60/292,335
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/297,523
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,925
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,810
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/303,807
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/303,808
```

```
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/315,047
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: US 60/324,928
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US 60/330,867
;  PRIOR FILING DATE: 2001-11-01
;  PRIOR APPLICATION NUMBER: US 60/330,462
;  PRIOR FILING DATE: 2001-10-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2221
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1412
;   LENGTH: 1310
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. D29960
US-10-152-319A-1412

Alignment Scores:
Pred. No.:              3.57e-35      Length:         1310
Score:                  336.00        Matches:        71
Percent Similarity:     57.14%        Conservative:   29
Best Local Similarity:  40.57%        Mismatches:     53
Query Match:            36.68%        Indels:         22
DB:                     17            Gaps:           3

US-10-657-740-1 (1-173) x US-10-152-319A-1412 (1-1310)

Qy           3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
               |||   :::|||      |||   :::|||          |||                
Db          18 GTGCCTGTGCAGCCTTCTTGGCTGCGCCGTGCTTCAGCTCCTTTACCGGGTTTTTCCACT 77

Qy          19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
               |||   ||||||||||||||   ||||||||||||||   |||   :::|||        
Db          78 CCGGGACGCCTCTTTGACCAGCGTTTCGGCGAAGGGCTGCTTGAGGCAGAGCTGGCTTCA 137

Qy          39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                     :::   |||:::|||||||||          |||           :::       
Db         138 CTGTGCCCTGCTGCGATCGCGCCCTACTAT---------CTGCGCGCCCCCAGTGTGGCG 188

Qy          59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
                     ::::::|||   :::|||          |||   :::   ||||||||||||    
Db         189 TTACCCACAGCCCAGGTGCCCACGGACCCTGGGTATTTTTCTGTGCTGCTGGATGTGAAG 248

Qy          79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
               |||||||||||||||:::::::::|||||||||      |||   ||||||:::|||    
Db         249 CACTTCTCGCCAGAGGAAATCTCTGTCAAGGTGGTTGGTGACCATGTGGAGGTCCATGCT 308

Qy          99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
               :::|||   ||||||   |||:::||||||:::|||:::|||||||||||||||||||||
Db         309 CGGCATGAGGAGCGCCCAGATGAACATGGATTCATTGCTCGAGAGTTCCACCGCCGATAC 368

Qy         119 ArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMet 138
               |||||||||      |||||   :::|||::::::   :::||||||   :::|||:::  
Db         369 CGCCTGCCTCCTGGCGTGGACCCTGCTGCAGTGACCTCTGCACTGTCTCCTGAGGGTGTT 428

Qy         139 LeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAla 158
               |||:::                                                :::|||
Db         429 CTGTCTATC----------------------------------------------CAGGCC 443

Qy         159 IleProValSerArgGluGluLysProThrSerAlaProSerSer 173
                     |||   |||   :::             |||   |||::::::            
Db         444 ACACCAGCGTCGGCCCAGGCCTCACTTCCGTCACCACCTGCTGCC 488

RESULT 38
US-10-425-115-173881
; Sequence 173881, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 173881
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_9015C.1
US-10-425-115-173881

Alignment Scores:
Pred. No.:              1.05e-35      Length:         381
Score:                  334.00        Matches:        65
Percent Similarity:     74.14%        Conservative:   21
Best Local Similarity:  56.03%        Mismatches:     24
Query Match:            36.46%        Indels:         6
DB:                     18            Gaps:           4

US-10-657-740-1 (1-173) x US-10-425-115-173881 (1-381)

Qy           1 MetAspValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPheTyr---Pro 19
               ||||||:::   |||   |||||||||   :::|||         ||||||:::   |||
Db          35 ATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTTTCCACTCCCCC 94

Qy          20 SerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPhe 39
               ||||||||||||||||||||||||||||||||||   |||   |||   ||||||   |||
Db          95 AGCCGCCTCTTTGACCAGTTCTTCGGAGAGCACCTGTTGGAGTCTGATCTTTTCCCG--- 151

Qy          40 LeuSerSerThrIleSerProTyrTyrArgGln------SerLeuPheArg------Thr 55
                  |||::::::::::||||||:::|||   :::       |||       |||       :::
Db         152 ACGTCTACTTCCCTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGC 211

Qy          56 ValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeu 75
                     |||:::|||:::|||||||:::|||   :::::::|||:::|||   :::   |||
Db         212 TGGTTTGACACTGGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTG 271

Qy          76 AspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 95
               |||||||||||||||||||||||||:::|||   |||||||||      |||   :::|||
Db         272 GATGTGAAGCACTTCTCCCCAGAGGAACTCAGAGTTAAGGTGTTGGGAGATGTGATTGAG 331

Qy          96 IleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSer 111
               :::||||||||||||   ||||||||||||:::||||||:::||||||
Db         332 GTGCATGGAAAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCC 379

RESULT 39
US-09-960-352-4143
; Sequence 4143, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4143
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 18-LIB3058-040-Q1-K1-E5
US-09-960-352-4143

Alignment Scores:
```

```
Pred. No.:              1.67e-35        Length:         380
Score:                  332.50          Matches:        65
Percent Similarity:     75.65%          Conservative:   22
Best Local Similarity:  56.52%          Mismatches:     23
Query Match:            36.30%          Indels:         5
DB:                     9               Gaps:           2

US-10-657-740-1 (1-173) x US-09-960-352-4143 (1-380)

Qy         54 ArgThrValLeuAsp---SerGlyIleSerGluValArgSerAspArgAspLysPheVal 72
              ||||||   ||||||   |||      ||||||:::|||   ::::::|||:::|||
Db         12 CGCACCCAGCTGGATTGCTCTGCCCCCTCAGAGATGCGCCTGGAGAAGGACAGATTCTCT 71

Qy         73 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 92
              :::   ||||||:::   |||||||||||||||:::|||      ||||||      |||
Db         72 GTCAACCTGGACCTGACGCACTTCTCCCCAGAGGAACTCAAGGCCAAGGTGCTGGGAGAT 131

Qy         93 PheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArg 112
              :::|||:::|||||||||||||   |||||||||||:::|||||:::||||||||||
Db        132 GTGATTGAGGTGCATGGCAAACATGAAGAGCGCCAGGATGAACATGGTTTTATCTCCCGG 191

Qy        113 GluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 132
              ||||||||||:::||||||:::|||::::::||||||      |||::::::   |||
Db        192 GAGTTCCACAGGAAATACCGGATCCCAGCTGACGTGGACCCTCTCGCCATTACTTCATCC 251

Qy        133 LeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAla 152
              ||||||:::||||||   ||||||      ||||||:::   |||            |||
Db        252 CTGTCCTCTGATGGGGACCTCACTGTGAATGGACCAAGGAAACAG-----------GCC 299

Qy        153 ThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 167
              :::      ||||||   ||||||::::::|||||||||||||||
Db        300 TCCGGCCCTGAGCGCACCATTCCCATAACCCGTGAAGAGAAGCCG 344

RESULT 40
US-09-918-995-4075
; Sequence 4075, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4075
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-4075

Alignment Scores:
Pred. No.:              1.38e-32        Length:         405
Score:                  311.50          Matches:        60
Percent Similarity:     63.64%          Conservative:   24
Best Local Similarity:  45.45%          Mismatches:     41
Query Match:            34.01%          Indels:         7
DB:                     10              Gaps:           2

US-10-657-740-1 (1-173) x US-09-918-995-4075 (1-405)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
              |||   :::|||      |||   :::|||         |||
Db         16 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 75

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db         76 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 135

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
```

```
                 :::|||::::::|||||||||         |||             :::
Db        136 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 186

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
              :::::::::|||   :::|||         |||   :::   ||||||||||||
Db        187 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 246

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        247 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 306

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
              :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||
Db        307 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 366

Qy        119 ArgLeuProSerAsnValAspGlnSerAlaLeuSer 130
              |||||||||      ||||||   :::|||::::::
Db        367 CGCCTGCCGCCTGGCGTGGATCCGGCTGCCGTGACG 402

RESULT 41
US-10-723-860-380
; Sequence 380, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 380
;   LENGTH: 387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-380

Alignment Scores:
Pred. No.:              1.46e-30        Length:         387
Score:                  296.50          Matches:        57
Percent Similarity:     63.41%          Conservative:   21
Best Local Similarity:  46.34%          Mismatches:     38
Query Match:            32.37%          Indels:         7
DB:                     18              Gaps:           2

US-10-657-740-1 (1-173) x US-10-723-860-380 (1-387)

Qy          3 ValThrIleGlnHisProTrpPheLysArgThrLeuGlyProPhe-----------Tyr 18
              |||   :::|||      |||   :::|||         |||
Db         25 GTGCCTGTGCAGCCGTCTTGGCTGCGCCGCGCCTCGGCCCCGTTGCCCGGACTTTCGGCG 84

Qy         19 ProSerArgLeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuPro 38
              |||   ||||||||||||||   ||||||||||||||   |||   :::|||
Db         85 CCCGGACGCCTCTTTGACCAGCGCTTCGGCGAGGGGCTGCTGGAGGCCGAGCTGGCTGCG 144

Qy         39 PheLeuSerSerThrIleSerProTyrTyrArgGlnSerLeuPheArgThrValLeuAsp 58
                 :::|||::::::|||||||||         |||             :::
Db        145 CTCTGCCCCACCACGCTCGCCCCCTACTAC---------CTGCGCGCACCCAGCGTGGCG 195

Qy         59 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 78
              :::::::::|||   :::|||         |||   :::   ||||||||||||
Db        196 CTGCCCGTCGCCCAGGTGCCGACGGACCCCGGCCACTTTTCGGTGCTGCTAGACGTGAAG 255

Qy         79 HisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGly 98
              ||||||||||||||::::::   |||||||||      :::   ||||||:::|||
Db        256 CACTTCTCGCCGGAGGAAATTGCTGTCAAGGTGGTGGGCGAACACGTGGAGGTGCACGCG 315

Qy         99 LysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyr 118
```

```
            :::|||   ||||||   |||:::||||||:::::::::|||||||||||||||||||
Db        316 CGCCACGAGGAGCGCCCGGATGAGCACGGATTCGTCGCGCGCGAGTTCCACCGTCGCTAC 375

Qy        119 ArgLeuPro 121
              |||||||||
Db        376 CGCCTGCCG 384
```

RESULT 42
US-10-621-901-2279
; Sequence 2279, Application US/10621901
; Publication No. US20040067516A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisnewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-8-C3
; CURRENT APPLICATION NUMBER: US/10/621,901
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/319,414
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 2313
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2279
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-10-621-901-2279

Alignment Scores:

Pred. No.:	3.34e-30	Length:	543
Score:	295.50	Matches:	64
Percent Similarity:	60.14%	Conservative:	22
Best Local Similarity:	44.76%	Mismatches:	50
Query Match:	32.26%	Indels:	7
DB:	17	Gaps:	2

US-10-657-740-1 (1-173) x US-10-621-901-2279 (1-543)

```
Qy          7 HisProTrpPheLysArgThrLeuGlyProPheTyrProSerArgLeuPheAspGlnPhe 26
              |||   |||::::::                  :::|||   ||||||   ||||||
Db        120 CACGACTGGTGGGAG---------GATTACGACCACCCAATGCGTCTTATGGACCAACAT 170

Qy         27 PheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThrIleSer--- 45
              ||||||   ||||||          ||||||:::      |||   :::|||   |||
Db        171 TTTGGAATGGGCCTAAATCGTAATGATCTTATTACTAATCTAAGGGCCACTCCATCACTC 230

Qy         46 ---------ProTyrTyrArgGlnSerLeuPheArgThrValLeuAspSerGlyIleSer 62
                      |||||||||                     |||         |||
Db        231 TTCCGTGGCGGTTATTACAGACCTTGGAGGAATGAAATTACTGCTGACGATTCTTCATCA 290

Qy         63 GluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerPro 82
                 :::   :::|||:::|||||||||   :::   |||||||||:::   |||   |||
Db        291 ACTATCGTTGCTGACAAAGATAAATTCCAAGTGACTTTAGACGTTCAACAATTCAAACCA 350

Qy         83 GluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGlu 102
              :::::::::|||||||||   :::|||:::   |||   :::   |||||||||   |||
Db        351 AAAGAAATCACCGTAAAAACAAAGGACAATTGCGTAATCGTCGAAGGCAAACACGAAGAA 410

Qy        103 ArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSer 122
              :::||||||:::|||||||||||||||||||   |||   ||||||||   ||||||
Db        411 AAACAAGATGAACACGGATACATTTCCCGTCATTTTGTTCGACGATATGTCTTGCCCAGA 470

Qy        123 AsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCys 142
              |||   |||   :::   :::      |||||||||:::||||||:::||||||
Db        471 AATCACGATGCTGCCGATGTAGTATCGAGTTTGTCCTCGGATGGAGTATTGACCATAACA 530

Qy        143 GlyProLys 145
              ||||||
Db        531 GCGCCCAAG 539
```

RESULT 43
US-09-990-747-28
; Sequence 28, Application US/09990747
; Publication No. US20020081688A1
; GENERAL INFORMATION:
; APPLICANT: Kamb et al.
; TITLE OF INVENTION: Retinoid Pathway Assays, and Compositions Therefrom
; FILE REFERENCE: 29345/36934A
; CURRENT APPLICATION NUMBER: US/09/990,747
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/249,468
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 08/812,994
; PRIOR FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Vaccinia virus
US-09-990-747-28

Alignment Scores:

Pred. No.:	7.1e-30	Length:	349
Score:	291.00	Matches:	57
Percent Similarity:	76.42%	Conservative:	24
Best Local Similarity:	53.77%	Mismatches:	19
Query Match:	31.77%	Indels:	6
DB:	9	Gaps:	4

US-10-657-740-1 (1-173) x US-09-990-747-28 (1-349)

```
Qy         24 AspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeuSerSerThr 43
              |||   :::|||   |||   :::   |||   ||||||   |||      |||::::::
Db         24 GACGTCTACTTCCCTGAGTCCGTG---GAGTCTGATCTTTTCCCG---ACGTCTACTTCC 77

Qy         44 IleSerProTyrTyrArgGln------SerLeuPheArg------ThrValLeuAspSer 59
              :::||||||:::|||   :::      |||      |||      :::      |||:::
Db         78 CTGAGTCCCTTCTACCTTCGGCCACCCTCCTTCCTGCGGGCACCCAGCTGGTTTGACACT 137

Qy         60 GlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHis 79
              |||:::||||||:::|||   :::::::|||:::|||   :::   ||||||||||||||
Db        138 GGACTCTCAGAGATGCGCCTGGAGAAGGACAGGTTCTCTGTCAACCTGGATGTGAAGCAC 197

Qy         80 PheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLys 99
              ||||||||||||:::|||   |||||||||      |||   :::|||:::||||||||
Db        198 TTCTCCCCAGAGGAACTCAAAGTTAAGGTGTTGGGAGATGTGATTGAGGTGCATGGAAAA 257

Qy        100 HisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArg 119
              |||   |||||||||||:::||||||:::||||||||||||||||||||:::||||||
Db        258 CATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCCACAGGAAATACCGG 317

Qy        120 LeuProSerAsnValAsp 125
              :::|||::::::||||||
Db        318 ATCCCAGCTGATGTAGAC 335
```

RESULT 44
US-10-605-498-91
; Sequence 91, Application US/10605498
; Publication No. US20040127441A1
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rocchi, Palma
; APPLICANT: Signaevsky, Maxim
; TITLE OF INVENTION: Compositions and Methods for Treatment of Prostate and Other
; TITLE OF INVENTION: Cancers
; FILE REFERENCE: UBC.P-031
; CURRENT APPLICATION NUMBER: US/10/605,498
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US 60/415,859
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/463,952
; PRIOR FILING DATE: 2003-04-18

```
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91
;   LENGTH: 764
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-605-498-91

Alignment Scores:
Pred. No.:              3.7e-29       Length:         764
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     18            Gaps:           6

US-10-657-740-1 (1-173) x US-10-605-498-91 (1-764)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
              |||          ||||||                    :::|||||   ||||||
Db         47 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 106

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
              ||||||||||||   ||||||                        ||||||   |||
Db        107 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 142

Qy         40 ------------------------------------------------------------ 40

Db        143 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 202

Qy         41 -----------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                              :::      :::|||   |||   :::::|||   |||
Db        203 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 262

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
                 |||   ||||||:::||||||:::|||         |||:::::   :::
Db        263 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 319

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
              |||||||||   ||||||:::|||::::::||||||||||||   :::|||     |||
Db        320 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 379

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
              ||||||   |||||||||   ||||||||||||:::||||||||||||||||   |||
Db        380 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 439

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
                 |||:::|||   ||||||       ||||||   :::   :::|||   |||||||||
Db        440 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 499

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
                ::::|||   ||||||      ||||||                          |||
Db        500 CCTGAGGGCACACTGACCGTGGA-GGCCCC-------------------------CAT 531

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
              |||:::   :::      |||   |||:::      |||:::
Db        532 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 573

RESULT 45
US-10-641-643-1300
; Sequence 1300, Application US/10641643
; Publication No. US20040077003A1
;   GENERAL INFORMATION:
;        APPLICANT: Cocks, Benjamin G.
;                   Susan G. Stuart
;                   Jeffrey J. Seilhamer
;        TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;                            GENE EXPRESSION
;        NUMBER OF SEQUENCES: 1508
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;             STREET: 3174 PORTER DRIVE
;             CITY: PALO ALTO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/641,643
;             FILING DATE: 14-Aug-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: <Unknown>
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Zeller, Karen J.
;             REGISTRATION NUMBER: 37,071
;             REFERENCE/DOCKET NUMBER: PA-0001 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650) 855-0555
;             TELEFAX: (650) 845-4166
;    INFORMATION FOR SEQ ID NO: 1300:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 789 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GENBANK
;             CLONE: g32477
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1300 :
US-10-641-643-1300

Alignment Scores:
Pred. No.:              3.88e-29      Length:         789
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     17            Gaps:           6

US-10-657-740-1 (1-173) x US-10-641-643-1300 (1-789)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
              |||          ||||||                    :::|||||   ||||||
Db         74 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 133

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
              ||||||||||||   ||||||                        ||||||   |||
Db        134 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 169

Qy         40 ------------------------------------------------------------ 40

Db        170 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 229

Qy         41 -----------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                              :::      :::|||   |||   :::::|||   |||
Db        230 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 289

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
                 |||   ||||||:::||||||:::|||         |||:::::   :::
Db        290 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 346

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
              |||||||||   ||||||:::|||::::::||||||||||||   :::|||     |||
Db        347 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 406

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
              ||||||   |||||||||   ||||||||||||:::||||||||||||||||   |||
Db        407 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 466

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
```

```
Db        467 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 526

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
Db        527 CCTGAGGGCACACTGACCGTGGA-GGCCCC--------------------------CAT 558

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
Db        559 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 600

RESULT 46
US-10-153-668-283
; Sequence 283, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
;  APPLICANT: HONDA, Goichi
;  APPLICANT:  MATSUDA, Akio
;  APPLICANT:  MURAMATSU, Shuji
;  APPLICANT:  ISHIZAWA, Kenya
;  TITLE OF INVENTION: STAT6 Activating Gene
;  FILE REFERENCE: 1254-0207P
;  CURRENT APPLICATION NUMBER: US/10/153,668
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/293,172
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/316,031
;  PRIOR FILING DATE: 2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/328,403
;  PRIOR FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: JP 2001-157043
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: JP 2001-260681
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: JP 2001-313175
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 488
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 283
;   LENGTH: 847
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (108)..(722)
US-10-153-668-283

Alignment Scores:
Pred. No.:              4.32e-29      Length:         847
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     14            Gaps:           6

US-10-657-740-1 (1-173) x US-10-153-668-283 (1-847)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
Db        129 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 188

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
Db        189 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 224

Qy         40 ------------------------------------------------------------ 40

Db        225 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 284

Qy         41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
Db        285 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 344

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
Db        345 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 401

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
Db        402 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 461

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
Db        462 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 521

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
Db        522 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 581

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
Db        582 CCTGAGGGCACACTGACCGTGGA-GGCCCC--------------------------CAT 613

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
Db        614 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 655

RESULT 47
US-09-969-034-4480
; Sequence 4480, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
;  APPLICANT: Burgess, Christopher C.
;  APPLICANT:  Astle, Jon H.
;  APPLICANT:  Carroll, Eddie III
;  APPLICANT:  Catino, Theodore J.
;  APPLICANT:  Dwivedi, Poornima
;  APPLICANT:  Molino, Gary A.
;  APPLICANT:  Thiagalingam, Arunthathi
;  APPLICANT:  Lewis, Marcia E.
;  TITLE OF INVENTION: Nucleic Acid Sequences Differentially
;  TITLE OF INVENTION:  Expressed in Cancer Tissue
;  FILE REFERENCE: 1657/1032
;  CURRENT APPLICATION NUMBER: US/09/969,034
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: 60/237,271
;  PRIOR FILING DATE: 2000-02-10
;  NUMBER OF SEQ ID NOS: 4494
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4480
;   LENGTH: 865
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-034-4480

Alignment Scores:
Pred. No.:              4.45e-29      Length:         865
Score:                  289.50        Matches:        75
Percent Similarity:     50.52%        Conservative:   23
Best Local Similarity:  38.66%        Mismatches:     43
Query Match:            31.60%        Indels:         54
DB:                     11            Gaps:           6

US-10-657-740-1 (1-173) x US-09-969-034-4480 (1-865)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
Db        129 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 188

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
Db        189 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 224

Qy         40 ------------------------------------------------------------ 40

Db        225 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 284
```

```
Qy          41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                                  :::     :::|||   |||   ::::::|||   |||
Db         285 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 344

Qy          55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
                  |||   |||||:::||||||:::|||         |||::::::   :::
Db         345 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 401

Qy          75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
               |||||||||   ||||||:::|||::::::||||||||||||   :::|||      |||
Db         402 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 461

Qy          95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
               ||||||   |||||||||   ||||||||||||:::||||||||||||||||||   |||
Db         462 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 521

Qy         115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
                  |||:::|||   ||||||      ||||||   :::   :::|||   |||||||||
Db         522 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 581

Qy         135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
                  :::|||   ||||||   ||||||                              |||
Db         582 CCTGAGGGCACACTGACCGTGGA-GGCCCC--------------------------CAT 613

Qy         155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
               |||:::   :::      |||   |||:::      |||:::
Db         614 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 655
```

RESULT 48
US-10-172-118-626
; Sequence 626, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 626
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_001540
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-626

Alignment Scores:

Pred. No.:	4.45e-29	Length:	865
Score:	289.50	Matches:	75
Percent Similarity:	50.52%	Conservative:	23
Best Local Similarity:	38.66%	Mismatches:	43
Query Match:	31.60%	Indels:	54
DB:	17	Gaps:	6

US-10-657-740-1 (1-173) x US-10-172-118-626 (1-865)

```
Qy          10 PheLysArgThrLeuGlyPro---------------------PheTyrPro---SerArg 21
                  |||         ||||||                     :::||||||   ||||||
Db         129 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 188

Qy          22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
               ||||||||||||   ||||||                         ||||||   |||
Db         189 CTCTTCGACCAGGCCTTCGGG------------------------CTGCCCCGGCTGCCG 224

Qy          40 ------------------------------------------------------------ 40

Db         225 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 284

Qy          41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                                  :::     :::|||   |||   ::::::|||   |||
Db         285 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 344

Qy          55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
                  |||   |||||:::||||||:::|||         |||::::::   :::
Db         345 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 401

Qy          75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
               |||||||||   ||||||:::|||::::::||||||||||||   :::|||      |||
Db         402 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 461

Qy          95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
               ||||||   |||||||||   ||||||||||||:::||||||||||||||||||   |||
Db         462 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 521

Qy         115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
                  |||:::|||   ||||||      ||||||   :::   :::|||   |||||||||
Db         522 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 581

Qy         135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
                  :::|||   ||||||   ||||||                              |||
Db         582 CCTGAGGGCACACTGACCGTGGA-GGCCCC--------------------------CAT 613

Qy         155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
               |||:::   :::      |||   |||:::      |||:::
Db         614 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 655
```

RESULT 49
US-10-342-887-626
; Sequence 626, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 626
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-626

Alignment Scores:

Pred. No.:	4.45e-29	Length:	865
Score:	289.50	Matches:	75
Percent Similarity:	50.52%	Conservative:	23
Best Local Similarity:	38.66%	Mismatches:	43
Query Match:	31.60%	Indels:	54
DB:	17	Gaps:	6

```
US-10-657-740-1 (1-173) x US-10-342-887-626 (1-865)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
              |||             ||||||                        :::||||||   ||||||
Db        129 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 188

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
              ||||||||||||   ||||||                    ||||||   |||
Db        189 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 224

Qy         40 ------------------------------------------------------------ 40

Db        225 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 284

Qy         41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                              :::      :::|||   |||   ::::::|||   |||
Db        285 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 344

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
                 |||   ||||||:::||||||:::|||          |||:::::   :::
Db        345 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 401

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
              |||||||||   ||||||:::|||::::::||||||||||||   :::|||      |||
Db        402 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 461

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
              ||||||   |||||||||   ||||||||||||:::||||||||||||||||||   |||
Db        462 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 521

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
                 |||:::|||   ||||||       ||||||   :::   :::|||   |||||||||
Db        522 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 581

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
                 :::|||   ||||||      ||||||                          |||
Db        582 CCTGAGGGCACACTGACCGTGGA-GGCCCC-------------------------CAT 613

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
              |||:::   :::      |||   |||:::      |||:::
Db        614 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 655

RESULT 50
US-09-880-107-3865
; Sequence 3865, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3865
;   LENGTH: 1231
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z23090
US-09-880-107-3865

Alignment Scores:
Pred. No.:              7.57e-29        Length:         1231
Score:                  289.50          Matches:        75
Percent Similarity:     50.52%          Conservative:   23
Best Local Similarity:  38.66%          Mismatches:     43
Query Match:            31.60%          Indels:         54
DB:                     9               Gaps:           6

US-10-657-740-1 (1-173) x US-09-880-107-3865 (1-1231)

Qy         10 PheLysArgThrLeuGlyPro--------------------PheTyrPro---SerArg 21
              |||             ||||||                        :::||||||   ||||||
Db        513 TTCTCGCTCCTGCGGGGCCCCAGCTGGGACCCCTTCCGCGACTGGTACCCGCATAGCCGC 572

Qy         22 LeuPheAspGlnPhePheGlyGluGlyLeuPheGluTyrAspLeuLeuProPheLeu--- 40
              ||||||||||||   ||||||                    ||||||   |||
Db        573 CTCTTCGACCAGGCCTTCGGG----------------------CTGCCCCGGCTGCCG 608

Qy         40 ------------------------------------------------------------ 40

Db        609 GAGGAGTGGTCGCAGTGGTTAGGCGGCAGCAGCTGGCCAGGCTACGTGCGCCCCCTGCCC 668

Qy         41 ------------------SerSerThrIleSerProTyrTyrArgGlnSerLeuPheArg 54
                              :::      :::|||   |||   ::::::|||   |||
Db        669 CCCGCCGCCATCGAGAGCCCCGCAGTGGCCGCGCCCGCCTACAGCCGCGCGCTCAGCCGG 728

Qy         55 ThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePhe 74
                 |||   ||||||:::||||||:::|||          |||:::::   :::
Db        729 CAA---CTCAGCAGCGGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGCGTGTCC 785

Qy         75 LeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheVal 94
              |||||||||   ||||||:::|||::::::||||||||||||   :::|||      |||
Db        786 CTGGATGTCAACCACTTCGCCCCGGACGAGCTGACGGTCAAGACCAAGGATGGCGTGGTG 845

Qy         95 GluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPhe 114
              ||||||   |||||||||   ||||||||||||:::||||||||||||||||||   |||
Db        846 GAGATCACCGGCAAGCACGAGGAGCGGCAGGACGAGCATGGCTACATCTCCCGGTGCTTC 905

Qy        115 HisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 134
                 |||:::|||   ||||||       ||||||   :::   :::|||   |||||||||
Db        906 ACGCGGAAATACACGCTGCCCCCCGGTGTGGACCCCACCCAAGTTTCCTCCTCCCTGTCC 965

Qy        135 AlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHis 154
                 :::|||   ||||||      ||||||                          |||
Db        966 CCTGAGGGCACACTGACCGTGGA-GGCCCC-------------------------CAT 997

Qy        155 AlaGluArgAlaIleProValSerArgGluGluLysProThr 168
              |||:::   :::      |||   |||:::      |||:::
Db        998 GCCCAAGCTAGCCACGCAGTCCAACGAGATCACCATCCCAGT 1039

Search completed: May 30, 2005, 09:27:34
Job time : 1225.26 secs
```